INTEGRATING GMB AND GAMES IN LONDON'S BUILT ENVIRONMENT

By Shane Alan Carnohan

Supervised by: Professor Etiënne Rouwette, *Radboud Universiteit*
Second Supervisor: Professor Pål Davidsen, *Universitetet i Bergen*

A thesis submitted to Nijmegen School of Management in partial fulfilment of the requirements for the degrees of:

M. Phil in System Dynamics (Universitetet i Bergen, Norway)
M. Sc. in System Dynamics (Universidade NOVA de Lisboa, Portugal)
M. Sc. in Business Administration (Radboud Universiteit Nijmegen, the Netherlands)

RADBOUD UNIVERSITY
August 2016

Radboud Universiteit Nijmegen

EMSD
European Master in System Dynamics

Acknowledgements

Increasingly, as a result of exposure to other ways of being, people in most parts of the world are able to reflect on their lives and exert agency in the hope of bringing about change. – Lock & Nguyen, 2010, p. 9

I came across this quote during an undergraduate anthropology course several years ago and it is a good description of the culturally diverse EMSD program. Over the past two years I have been forced to reflect deeply on my own life and culture, and as a result, I developed a passion to create positive societal and personal changes. I now have the necessary capacities to achieve this thanks to the skills and knowledge gained in this program, and for that, I express my gratitude to the entirety of the program, including its founders, professors, and administrators.

I must also express thanks to the mentors and advisers that guided me through the completion of this thesis: To Dr. Nicole Zimmermann, my host at UCL whose modeling skill and intellect is matched only by her ability to defend students from building security. To Dr. Etienne Rouwette, whose courses in group model-building built inspired me pursue this participatory work. And to Dr. Pål Davidsen, who gave me the opportunity to see the Midnight Sun and hike among the reindeer.

The journey would have lost all of its richness without my EMSD 'family.' We shared many unforgettable memories and became lifelong friends in the process.

A special thanks is owed to my friends in Costa and Hoedspruit who made nice settings for thesis writing into paradises of productivity. Thanks goes to also to my London 'roommates' who taught me how to survive the urban jungle, and reminded me not to take life too seriously.

I will always be indebted to Dr. Andrew Ford who first introduced me to the virtuous system dynamics lens. Those early lessons continue to guide my thinking today. I am also grateful to Jake Jacobson for providing friendship, encouragement and good coffee.

Finally, no words can adequately describe my appreciation for the support of my family. To my father, for the gifts of pragmatism and honesty. To my mother Lisa, for teaching me the importance of kindness to others as well as one's self. And to my big brother, for showing me the importance of acceptance and always leading so I could follow. The completion of this thesis journey is as much their accomplishment as it is my own.

Abstract:

A participatory research process was carried out with stakeholders in the domain of the built environment in London, U.K. The objective of the study was to improve stakeholder capacity for integrated decision-making by addressing multiple objectives of the built environment while examining the relative contributions of group model building (GMB) and simulation games to group processes. This was done in order to reduce fragmentation, or a lack of integrated planning, among London's built environment decision makers, and to add to the understanding of how system dynamics-based simulation environments or games can be used effectively in participatory GMB process. Therefore, GMB and a simulation game were applied in an integrated process and outcomes were assessed on the basis of questionnaires, observational data and audio recordings of the sessions. The integrated process lead to improvements in participant learning, and developed shared understandings among stakeholders. This is evidence that the process was successful in reducing fragmentation. In addition, scales measuring learning and commitment were found to be higher in the game workshops than in GMB workshops, which were evaluated more positively on scales for consensus and communication. These differences are interpreted on the basis of transcribed audio data. An overall small sample size and other difficulties reduced the reliability of the results. However, the novel aspects of this design provide encouraging implications for future research regarding the contributions of games to facilitated group processes.

This research was funded in-part by UCL Innovation and Enterprise and the UK Engineering and Physical Sciences Research Council (EPSRC) Platform Grant EP/I02929X/1.

Table of Contents

List of Figures

List of Tables

Introduction

The built environment serves many roles in human society. It is where the increasingly urban human population lives, works and plays. It is made up of homes, offices, parks, pubs and the intervening elements in between. The built environment must simultaneously meet multiple, different goals, and therefore a more holistic understanding is needed in order to provide an environment where people can thrive, not just survive. In order to do so factors that make up social and individual wellbeing, which can be difficult to measure, must be addressed in a coordinated manner. Approaches are needed that can cope with these challenges in measurement, and investigate the interdependencies that make up the built environment. The design of future policies will benefit from more integrated planning that considers these interconnections, thereby enabling better performance within this complex system.

This thesis focuses on the city of London in the United Kingdom (U.K.) where aggressive policy targets for carbon emissions reductions has led to increased pressure on the housing sector to apply energy efficiency techniques (HM Government, 2011). These policies arose following the passage of the Climate Change Act, which calls for an 80% reduction from 1990 emissions levels throughout the U.K. (*Climate Change Act*, 2008, sec. c. 27). As housing emissions account for more than one quarter of total emissions, this sector has a large role to play in meeting these ambitious targets and more than 14 million homes in the U.K. are targeted for improvements in energy efficiency by 2020 (Department of Energy and Climate Change, 2012).

Thus far, U.K. housing policies have consistently underperformed, both in meeting their primary objective to reduce emissions contributed by the housing stock, and in mitigating unintended, unwanted consequences (Davies & Oreszczyn, 2012). This has been attributed to failures in policy development processes that have singular objectives, which has resulted in negative impacts on communities as well as the mental and physical wellbeing of residents (Shrubsole, Macmillan, Davies, & May, 2014). It has been suggested that, in order to improve the performance of policies in this complex domain, more holistic thinking must be combined with new methods that can better integrate multiple objectives into the planning process (Eker & Zimmermann, 2016; Eker, Zimmermann, & Carnohan, in preparation; Shrubsole et al., 2014; South, 2015). In practice, this requires decision makers to be engaged in processes that can develop their ability to deal with multiple policy goals successfully.

A project about Housing, Energy and Wellbeing (HEW) that addresses this gap has been underway at the University College London that focuses on integrated decision making. This work engaged stakeholders (who are subsequently engaged in this thesis research) using components of the system dynamics (SD) method. Specifically, this work used qualitative causal loop diagrams (CLDs), a tool applied as a part of the SD method. SD focuses on defining systems as cause and effect relationships between different elements, qualitatively or quantitatively, allowing the development of a more holistic view of the system. CLDs are a representation used in the SD method, that show the direction of these relationships (either positive or negative) and any time delays in between these in a qualitative manner. The SD approach emphasizes feedback, or how initial changes in one system element propagate through the

system over time, eventually returning back to effect a change in the initial element once more (Sterman, 2000).

The HEW project applied CLDs to address the complex challenges facing the built environment. Research began with stakeholder identification and individual interviews. Individual interviews were coded using an inductive process and organized into themes that showed the interconnections between social and technical factors of the built environment. The themes are a distilled representation highlighting the interconnections that emerged from the interviews (Macmillan et al., 2016). An understanding of these multiple dependencies is needed to improve performance of future policy designs and avoid unintended, unwanted outcomes (Shrubsole et al., 2014).

The themes were then refined and developed into nine policy criteria, shown in Table 1. The themes and resulting criteria collectively show the stakeholders' consideration of difficult to measure items such as community connection and other aspects influencing social wellbeing. It is notable that, besides the identification of these criteria, the second most discussed topic (behind energy efficiency) was social wellbeing. The interviews and themes were interpreted as representing improvement in shared understanding of the decisions made in the complex housing system among stakeholders (Macmillan et al., 2016). However, there is remaining need for further efforts toward integrative planning and consideration of the multiple objectives of the built environment pertaining to social, physical, and mental wellbeing.

This understanding was developed after a third workshop, following previous qualitative workshops, where stakeholders indicated that *fragmentation* or, in general, a lack of integrated planning, has led to noticeable gaps between intentions or planned designs and the implementation of these (Zimmermann, Black, Shrubsole, & Davies, 2015). This thesis is focused on tackling fragmentation that is present at the *individual level* and occurs between individual decision-makers. It also addresses intra-group fragmentation where implementation breaks down due to the lack of coordination among organizations. In addition it continues the use of the holistic SD approach, advancing beyond the CLD diagrams to address fragmentation and encourage stakeholder consideration of the impacts of policies on the previously described items. Further application of this approach has been suggested as a way to overcome fragmentation by enabling integrated planning and decision making activities (Eker & Zimmermann, 2016).

The methodology chosen is group model-building (GMB), which combines facilitated discussions with detailed action plans and diagramming conventions, like CLDs, to integrate stakeholder input (Andersen & Richardson, 1997; Vennix, 1996). This method has been shown to be effective for developing learning, building consensus, generating commitment and improving communication with client groups (Rouwette, Korzilius, Vennix, & Jacobs, 2011; Rouwette, Vennix, & Mullekom, 2002). It has also been demonstrated as useful for resolving management conflicts (Black & Andersen, 2012). In this study, GMB is combined with a simulation game, based on an SD-model. Like GMB, games can be an effective tool for participant learning (Davidsen & Spector, 2015). However, unlike GMB, little else is known about how these games influence different dimensions of intended group process outcomes.

*The objectives of the study are to improve stakeholder **capacity for integrated decision-making** by addressing the **multiple objectives** of the built environment and to examine the relative contributions of group model building and simulation games to group processes. This is done in order to reduce fragmentation among London's built environment decision makers and to add to the understanding of how simulation games can be used effectively in participatory GMB process.*

GMB and games have only been compared twice previously (Eskinasi & Rouwette, 2004; Ruud & Baakken, 2003), and only once on the basis of an established questionnaire(Eskinasi & Rouwette, 2004). This thesis added rigor by supplementing results from an established questionnaire with a thorough analysis of audio recordings as well as observational data. Analysis was also performed, on the basis of the audio recordings, in order to measure the extent to which these processes improve stakeholders' consideration of multiple objectives, such as wellbeing and community. Furthermore, this work builds on previous theory regarding the use of visual representations in group processes while piloting the use of a method for eliciting stakeholder knowledge during group process.

The remainder of this thesis is organized into a summary of the literature, followed by a detailed description of the small and large group workshops. The processes implemented and experienced by the facilitation team will then be described, summarizing the observational data. Next, the results are presented and interpreted, followed by a discussion of the limitations and a statement regarding ethical considerations. The thesis concludes by addressing areas for future research followed by the author's reflection, which contains insights about the conditions under which the research was conducted.

Chapter 1: Literature Review

This section begins with a summary of the HEW project background which establishes the complex challenges faced in the U.K. built environment and motivates the need for the participatory GMB approach. This is followed by an explanation of the GMB and game approaches along with a summary of relevant literature which compares the two methods in order to situate the contributions of this study to existing theory within the broader field.

Housing, Energy and Wellbeing (HEW) Project Background

This research extends from the qualitative application of system dynamics that has been conducted as part of the HEW project at the University College London. It recognizes housing as an area of baffling complexity. As a result it is an area where "policy resistance" is known to occur (Davies & Oreszczyn, 2012; Macmillan et al., 2016; Sterman, 2000). This term describes instances when "policies are delayed, diluted or defeated by the unforeseen reactions of other people or of nature"(Sterman, 2000, p. 3). These reactions have had negative consequences in other domains that are tightly coupled to the built environment. This is due to policy-makers' consideration of these multiple objectives in an isolated manner. The HEW project goals addressed this issue, as described by Macmillan et al. (2016):

> *"This research aimed to move from considering disparate objectives of housing policies in isolation to mapping the links between environmental, economic, social and health outcomes as a complex system. We aimed to support a broad range of housing policy stakeholders to improve their understanding of housing as a complex system through a collaborative learning process"* (2016 p. 1).

So far, the project has engaged a diverse group of more than 50 stakeholders, including government, academic, industry, community and non-governmental representatives. CLDs were developed through individual interviews. These were first distilled and the elaborated used in series of two collaborative causal mapping workshops. The resulting diagrams were reported as useful in facilitating learning among the stakeholder group, based on results of coded interviews, and were later used to inform the identification of the stakeholder group's top nine policy criteria, shown in Table 1 below. These placed emphasis on technical aspects such as carbon emissions. However, they also emphasized less technical aspects such as community connection and factors that contribute to individual wellbeing (physical & mental health). Note in the table that community connection, a factor related to social wellbeing was ranked second among policy criterion. This makes sense, given that the second most discussed topic appearing in the code was social wellbeing. The diversity among these ranked criteria is evidence that the stakeholder group had developed a consideration of multiple objectives as a result of the project (Macmillan et al., 2016).

Table 1. The policy criteria developed by Macmillan et. al. (2016).

	Policy Criteria
1. *Carbon emissions from housing*	6. *Mental and emotional wellbeing*
2. *Community connection*	7. *Physical wellbeing/health*
3. *Fuel poverty*	8. *Policy Coherence*
4. *Housing adaptation to climate change*	9. *Social and Income Equality*
5. *Housing affordability*	

The work of Macmillan et al. (2016) established some shared understanding among the stakeholder group. However, during a third workshop a persistant problem surfaced, called *fragmentation* (Eker & Zimmermann, 2016). Follow-up interviews were then conducted with a smaller core group of the original HEW stakeholders which helped to define this phenomenon. In this study it is generally defined as absence of integrated planning by actors involved in carrying out policies in the built environment. More formally, can be considered as having three different levels: The first level of fragmentation takes place between individual decision-makers. The second level of fragmentation occurs between organizations or groups and the third level is the vertical divide between the higher level policy organizations and organizations at the local level.

The overall research process is shown in Figure 1, from the start of the project with stakeholder analysis and interviews, to this study's contribution of 3 GMB sessions plus the final game workshop. The stages are adapted from Macmillan et al. (2016), and the most recent steps, respond to the the authors' recommendations for the use of new approaches that can "integrate the qualitative and quantitative knowledge held by different groups[…] in a collaborative learning process[…] and explore the impacts of policies on a more integrated set of outcomes" (p. 2). In this study, the use of GMB to support the development of a quantified game represents a first iteration of this integration, and is therefore well situated to contribute to the overall HEW project goals.

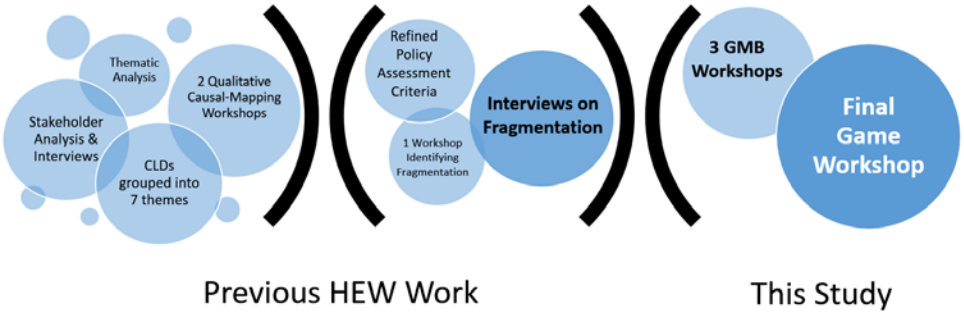


Figure 1. A flow-chart giving an overview of the HEW research program, developing from qualitative interviews towards this study's use of the GMB sessions and simulation game.

Fragmentation has been associated with the creation of gaps between intentions or planned designs and the implementation of these. In the built environment this manifests both as poor coordination among organizations and as narrowly focussed decision-making by individuals (Davies & Oreszczyn, 2012; Shrubsole et al., 2014). The influence of fragmentation on the social and individual factors of wellbeing and community are described next.

Communities and Wellbeing in the Built Environment

Other highly visible and influential entities, such as the World Health Organization (WHO), Public Health England and England's National Health Service (NHS), have independently arrived at findings similar to the MacMillan et al. (2016) study. This highlights the importance of this work and provides further support for the need to consider the wider influences the built environment has on social wellbeing (communities), and individual wellbeing (physical and mental health). These two general domains were chosen for emphasis in order to address the policy criteria determined by stakeholders previously. Next, a handful of studies are described that give some insight regarding the interplay between these factors and the built environment. This is done in order to illustrate the measurement difficulties that arise which motivates the application of the GMB method that is a useful tool for addressing subjective problem spaces such as these (Vennix, 1999).

Much of the literature available has attempted to break down the complexity into seperate components, in order to compare certain physical features of the built environment, such as green space, with self reported health measures. For example, the use of surveys or other techniques have been used to relate *physical* health to aspects such as walkability, level of crowding and green space (Francis, Giles-Corti, Wood, & Knuiman, 2012; Prochorskaite, Couch, Malys, & Maliene, 2016; Villanueva et al., 2015). A case study that took place in the highly relevant area of Greenwich, London was directed towards measuring the *mental* health of residents, as influenced by specific physical features and social aspects of the built environment. The study used different assessment scales in the form of questionnaires to gather data and found that perceived levels of noise, overcrowding, green spaces, safety, and community facilities were all significant predictors of mental health outcomes (Guite, Clark, & Ackrill, 2006). These results reinforce the links between the built environment and the wellbeing of residents, however the direction and strength of mechanisms by which these elements influence each other have not yet been well established.

These studies' inclusion of measurements such as noise levels and overcrowding exemplify a growing understanding that healthy individuals depend, at least partially, on a well-functioning *community*. Frumkin (2003) provides an excellent overview of the qualitative evidence that the design of the built environment has a large impact on human health and wellbeing. By covering psychological literature, observational research, architecture and design as well as some empirical health outcomes, they distill four categories within the built environment which impact wellbeing: buildings, public places, urban form and nature contact. These categories all help to define the "sense of place" for a community. This sentiment is shared by NHS and Public Health England. In a 2015 report, they indicate the importance of communities both as an operational level for implementing policy and as a socially constructed combination of knowledge and identity. They write about the "extensive evidence that connected and empowered communities are healthy communities" and conclude the following: "Communities that are

involved in decision-making about their area and the services within it [and], that are well networked ad supportive[…] have a positive impact on people's health and wellbeing. (South, 2015, p. 4)" These examples highlight the importance of built environment decision makers' consideration of the relationship between the life of the community and physical spaces that it inhabits, demonstrating the importance of this study's objectives.

In order to create such communities and improve overall resident wellbeing, the NHS recommends wider participation and integrated planning (South, 2015). This supports the need for the use of a participatory approach to tackle fragmentation which is seen to limit this integration. This is corroborated by the World Health Organization (2012). They recommend leveraging aspects of the built environment in order to support the creation of sustainable communities based on whole-system processes. Integrated plans can then be created in order to address the "wider social determinants of health " (p. 9) as well as to secure "stakeholder ownership" (p.9). These reports show the complexity and importance of the problem space addressed in this thesis, and lend support for the use of the participatory GMB method.

Situating Group Model-Building and "Games"

The complexity of the challenges within the U.K. built environment, coupled with the expressed need for a more participatory approach supports the use of the GMB method in the project. Stave (2010) gives several examples of how the method can be applied in practice to help decision makers deal with "messy" problems. Messy problems exist when there are "large differences of opinion on the problem or even on whether there is a problem, where there is no clear solution or perhaps no solution at all" (Stave, 2010, p. 2764). She highlights the decision support needs for sustained management of environmental systems, but the rationale can easily be applied to the built environment as well. For example, environmental decision makers face vast complexity, trade-offs and subjectively defined problems. The built environment, and the HEW project, are found within a similar context. Subjective judgment bleeds into the problem space, compounding complexity and rendering purely objective problem-solving measures inadequate.

Heeding the call for a participatory approach is the facilitated modeling approach, specifically GMB. Next, some methodological components of GMB will be explained, beginning with a description of system dynamics, the modeling method used to support these interventions. Afterwards, a description of SD-based games will be given, followed by a section which evaluates GMB and games along the dimensions that this study will use for comparison.

1. System Dynamics (Expert Mode)

SD is a methodology for the study of complex, non-linear systems over time (Ford, 2010; Forrester, 1971a, 1971b; Sterman, 2000). Pioneering work by Forrester (Forrester, 1961) established the basic theory, that explores socio-economic factors along with physical systems. These models provide a realistic approach to system simulations that incorporate "human error" elements such as informational delays and imperfect understanding. It is a method which analyzes the means and ends, or causal relationships, between different variables of a given system *over time*. Simulations can take the form of envisioned future scenarios where the established relationships between variables can give rise to the unexpected non-linear behavior of the system. Systems are rendered using three components, 1) stocks (where things *accumulate*), 2) flows (which determine how fast or slow stock values change) and 3) converters (defining causal relationships between variables). Models are typically created in a graphical user interface (GUI) and result in stock and flow (S&F) diagrams. High levels of transparency are enabled as S&F models, created using commercial software such as Vensim®, and Stella®, allow non-experts to engage with the model. This is sometimes referred to as *white-box* modeling, as opposed to *black-box* models that are hidden from view. This transparency can be one of the primary strengths of system dynamics over other modeling disciplines (Ford, 2010).

For this study it is also useful to consider the differences between 'expert' and 'facilitated' modes of SD modeling. Assumptions made by each mode are described by Franco & Montibeller (2010). They assert that the expert mode typically considers problems with an objective lens, and emphasizes the role of the modeler in defining the problem and arriving at optimal solutions. Any solutions given by the expert modeler, then, are assumed to result in commitment of the stakeholders, due to the objective nature of the analysis. Expert mode approaches are common in SD literature, however, the impacts of this mode on the client group remain unclear (Größler, 2007; Snabe & Größler, 2006). Of course, the two modes of expert and facilitated modeling represent extremes and plenty (if not most) SD modeling takes place in between these. (A classic example is Homer's (1985) worker burnout model, which is based on the author's experience with this issue.) However, applying the high-contrast expert/facilitated lens is useful for pointing out the benefits of the GMB approach in this review of the literature.

2. GMB (Facilitated Mode)

Group model-building is an effective facilitated modeling approach that incorporates SD and stakeholder input to increase learning and ownership (Rouwette et al., 2002). Facilitated modeling via GMB methods loosens the assumptions used in the expert modeling approach. Problems are considered to be subjective, rather than objective and clearly defined. This allows for introduction of metrics that are most useful for the client group. GMB encourages participation and assumes that involvement in the process will in fact create higher levels of commitment to the results (Franco & Montibeller, 2010). The effects of this mode on client groups has been extensively studied (Rouwette et al., 2002) and it is thought to increase the likelihood that the client or stakeholder group will act upon model findings (Vennix, 1999). The GMB method was defined by talented SD practitioners and academics who began to see new avenues for applying the method with client groups. The previously mentioned transparency of the SD method allows for ease of communication, thanks to the S&F and causal loop diagrams. Although such diagrammatic methods had existed for decades before the emergence of GMB, the arrival of GUI-

enabled SD software made this process much faster and easier (Hovmand, 2014). Previously, stencils would be used to hand-draw the various system icons and calculations took place using a code-based language (e.g. (Forrester, 1969)).

A helpful element in applying the facilitated approach are the GMB "scripts". These are detailed action plans which add structure to the group process, providing a baseline to make the workshop coherent while leaving room for unplanned discussions that add richness to the outcomes (Andersen & Richardson, 1997). Workshops are designed in phases where levels of beneficial divergent and convergent thinking are managed elements. Vennix (1996) describes these phases in terms of the levels of cognitive conflict occurring among the group. During a divergent phase, cognitive conflict is high, and individuals share multiple points of view to generate new ideas or new understanding. Then in convergent phases, cognitive conflict is declining and participants integrate the diverse perspectives into a new shared theory. Facilitated group processes in this style are helpful in overcoming some of the cognitive limitations individuals and groups experience such as overconfidence, anchoring and adjustment and group think (Tversky & Kahneman, 1974; Vennix, 1996).

3. SD Simulation Environments or "Games"

SD is a method with a long history of use in the design and use of games. In the literature, many different titles have been given to these SD-based simulation environments, such as "management flight simulators" or "interactive learning environments" (Alessi & Kopainsky, 2015; Andersen, Chung, Richardson, & Stewart, 1990; Davidsen & Spector, 2015; Ford, 1996; Kopainsky, Alessi, Pedercini, & Davidsen, 2015; Lane, 1995; Maier & Größler, 2000; Meadows, 1989; Ruth, 2015; Sterman, 1992; Van Daalen, Schaffernicht, & Mayer, 2014). Zimmermann et al. (2015) have applied an interactive approach within previous HEW-related work and discussed the implications for the use of games to create a shared understanding among stakeholders. This is not a new notion. Games are a means to "provoke, release and utilize personal and even emotional elements of learning (Lane, 1995, p. 612)" but can be facilitated in a way that adds an element of fun to the mix (Lane, 1995). The use of games to improve participant learning outcomes is summarized next.

A review by Davidsen & Spector (2015) summarizes recent contributions to the theory surrounding use and evaluation of games. They summarize recent contributions made to decision making, learning and policy development. With respect to learning, two different approaches dominate – inquiry learning and debriefing. Inquiry learning is demonstrated as an interface that is navigated by the user alone, without help from an outside facilitator. With this approach learning does take place and is related to the use of the interface, however the relative contributions of the different elements of this approach remain poorly understood (e.g. the participant's attitude at the start of the experiment) (K. A. Stave et al., 2014). Sterman et al. (2013) describe a good example of a SD-based Climate Rapid Overview And Decision Support (C-ROADS) model. It is a well-known and widely applied tool designed to "build shared understanding of climate dynamics (Sterman et al., 2012, p. 296)." It has been used at the UN Climate negotiations and in classrooms. A free version of the simulator, called C-LEARN is available for use on the web. However, this is primarily an inquiry learning experience and the authors indicate that experimentation and reflection aids the user to "learn for themselves" about climate change. The user

interface is intended to teach about the synergies between coordinated actions and emphasize importance of coordination among players (Sterman et al., 2012, 2013). It is also designed to allow modification of some assumptions, motivating more informed negotiations (Alessi & Kopainsky, 2015). However, the C-ROADS model is an example of a game that was developed in an expert fashion and the more detailed model underlying the simulations is not typically investigated in its use (Sterman et al., 2012). This can lead to participants rejecting the intended learning outcomes, based on the argument that alterations in the model's assumptions would result in different outcomes (Davidsen & Spector, 2015).

In contrast, the debriefing approach involves a facilitator who supports to the user of the game as complexity is gradually increased and explained. Kopainsky et al. (2015) refer to this as the *prior exploration strategy*. Prior exploration is seen as a distinct phase intended to overcome barriers to system management and it emphasizes the dynamic task and clarity of the user influence. Similarity between the prior exploration phase and subsequent *management* phases is necessary in order for participants to compare results. Finally, decisions should be reversible to overcome any participants' fear of uncertainty and risk. They applied these principles in a controlled experimental setting and found that engaging participants in this way led to a significant learning improvement (Kopainsky et al., 2015). The prior exploration approach bears similarity to the *trial-driven* process described by Van Daalen et al. (2014), where the model is initially hidden and becomes more transparent during a debrief and *analysis-driven* process (management phase). This two-stage approach has potential advantages as it can prevent participants from anchoring their ideas in the existing structure. It also ensures they are not overwhelmed by too much complexity. Furthermore, the cognitive dissonance created by the initial trial-driven simulation mode motivates system inquiry (Van Daalen et al., 2014). This attribute in particular, can be useful in GMB interventions as it can encourage participants to engage in the process. The integration of these two approaches is explored more in the following section.

Comparing GMB and Games

Both GMB and games have been determined to have positive effects on participant learning. In GMB this is attributed to elements of the *process* encountered while building a model with a small group of people. The elements often assessed are discussions, presence of a facilitator, use of diagrams (including CLDs) and simulations using the model (Rouwette et al., 2011). SD-based games have also been used to facilitate learning. However the assessment of learning using simulations and games has historically focused on the use of the modeling environment, rather than the process of playing the game (Davidsen & Spector, 2015). Evidence for the effectiveness of considering process along with use was given by (Kopainsky et al., 2015), when they applied the previously described prior exploration strategy.

The design and use of a game may also employ certain GMB elements. In the prior exploration strategy, for example, interaction with participants implicates facilitation as an important element. The debrief session that precedes the management phase can employ CLD diagrams or other visual representations to describe the feedback relationships that underlie the game behavior. However, the use of visual elements within each approach also differs. GMB has made extensive use of diagrams that serve to improve collaboration among participants, known as *boundary objects* (Star & Griesemer, 1989). Boundary objects are the "tangible representation of dependencies across disciplinary, organizational,

social or cultural lines that all participants can modify" (Black & Andersen, 2012, p. 195). Recent work demonstrates the way in which formal theory surrounding boundary objects can be related to GMB interventions. In this study where GMB and games are integrated, this was chosen as a means of analyzing whether or not the game acts to support this process and as well as a lens for interpretation of the observational data.

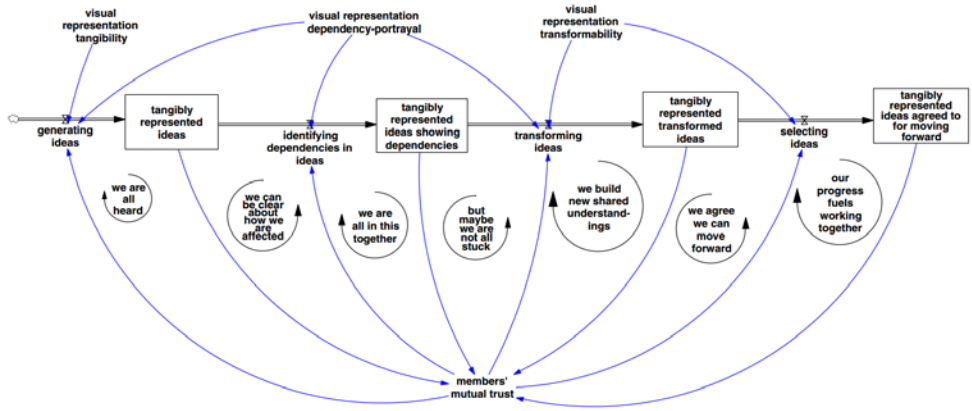


Figure 2. The facilitated process is supported by the use of boundary objects, which drive the accumulation of tangibly represented ideas and dependencies, transformed into ideas for moving forward (Black & Andersen, 2012).

The four stage process proposed by Black & Andersen (2012) is represented as a stock and flow diagram, shown in Figure 2, that accumulates understanding during workshop sessions. They also define three distinct characteristics of boundary objects, which must 1) be a tangible visual element, 2) show dependencies and 3) be modifiable by all participants. The integrated nature of the present study's research strategy disperses the phases to be captured in each workshop.

Games encourage critical analysis of the model structure and may indeed act as a boundary object. Some have asserted previously that it is a challenge to fulfill the boundary object requirement of transformability in a simulation setting (Black, 2013). However, others have suggested the use of games as boundary objects (Zimmermann et al., 2015). More investigation is needed to understand the theoretical basis which motivates the use of games.

Both GMB and games lack a standard evaluation method that can be used to relate intervention elements to outcomes (Davidsen & Spector, 2015; Rouwette et al., 2011, 2002). In addition, despite the increased emphasis on the process of game play towards achieving certain outcomes, such as learning, the two have rarely been compared. To better relate GMB process elements and their effects, Vennix et al. (1993, 2000) designed a questionnaire that introduced scales of consensus, insight, communication and commitment to action (CICC). This questionnaire has been shown as an effective way to add rigor to evaluation, serving as an example of a possible standard assessment tool for the method (Rouwette et al., 2011). Though effective measures on the basis of learning have been demonstrated in studies that use simulation environments, standard evaluation methods are also absent (Davidsen & Spector, 2015).

Therefore the use of a questionnaire may be one way to streamline data collection for both GMB and games, enabling a "more systematic assessment of projects and accumulation of research results (Rouwette et al., 2011, p. 886)." Additionally, it allows a comparison of the process elements which can help to further elucidate important elements of each approach.

Table 2. Dimensions of GMB and games that are compared in this thesis research.

Dimensions Compared	**GMB**	**Games**
Learning	✓ (Vennix, 1996)	✓ (Kopainsky et al., 2015)
Building consensus	✓ (Rouwette et al., 2011)	? (Ruud & Baakken, 2003)
Improving communication	✓ (Rouwette et al., 2011)	? (Ruud & Baakken, 2003)
Use of boundary objects	✓ (Black & Andersen, 2012)	? (Black, 2013; Zimmermann et al., 2015)

Despite the similarities between these two approaches, only three prior studies could be found that specifically combined GMB and games. The most recent study, relating to water and sustainable development, used one GMB session to create a CLD. The relationships defined in that session were used as input for the final model and game. The authors credit the GMB workshop for its contribution to the identification of key variables. However, they do not compare GMB to games, nor do they use any kind of systematic analysis to evaluate the specific contributions made by the GMB session (Bassi, Rego, Harrisson, & Lombardi, 2015).

Ruud & Baakken (2003) combined the methodologies to create a decision support tool for military training. They created a multiplayer game using GMB to inform the process and speak to the use of the approach for learning. They also point out "how people who have worked side by side for a long time could "update" their perception of each other's *understanding* during the modeling process [emphasis added]" (Ruud & Baakken, 2003, p. 6). As this process involved use of the gaming interface, their observation provides some evidence that games can be used to improve consensus and communication. However, this is weakly supported in the study. In addition, respondents in their study also "emphasized how the game is a tool for triggering group discussions "(Ruud & Baakken, 2003, p. 8). This implies the use of the game as a boundary object, however, beyond anecdotal and observational evidence this study provided little support for either.

Another study using both of these methods was also conducted in the realm of the built environment by Eskinasi & Rouwette (2004). Participants in their study used a 'flight simulator' for 15-30 minutes individually as part of a two-hour workshop presenting simulation runs to the larger group. This was part of a larger GMB case study (Eskinasi, Rouwette, & Vennix, 2009) that took place in the Netherlands, focusing on the tensions between new construction and the market for subsidized, social housing. They apply a pre-test, post-test design based on a measurement model of intended behaviors of participants. An example of a behavior in this case took the form of intended policies to address this tension. They also asked participants to compare the workshops to their experience of a normal meeting. They report that both groups found the workshops to be more effective than normal meetings. They also found a significant difference for two dimensions of behavior, but they do not provide any comparison of the

two on this basis. Taken together, these studies clearly illustrate that there are indeed theoretical gaps in understanding regarding the effects games on group processes. Furthermore it appears that there is perceived positive effect from combining games with GMB in this regard. Clearly, further investigation is still needed to understand how these two methods may complement each other.

Following the context laid out by this literature review, the research questions are defined. The methods section follows, and will expand upon this theoretical underpinning with a description of the research strategy. This will outline the way in which GMB and SD-games were integrated and how the effects were measured.

Research Questions

1. Can SD simulation games be integrated with GMB practices in a productive manner?
    a. Is one method more effective than the other at facilitating participant learning, consensus building, and communication?
        i. Is there a difference in performance for participants that participated in previous GMB work?
    b. Are games useful as boundary objects as a part of group process?
2. How does boundary object use compare between GMB and games?
3. Does the integrated process lead to an increased consideration of the multiple objectives of the built environment, specifically pertaining to social and individual wellbeing?
4. How can weighting techniques be used in GMB sessions to elicit participant values for some represented model variables constructively?

Chapter 2: Methodology

The research strategy is a case study employing mixed methods of data collection. A facilitated modeling approach is chosen as the primary method of problem structuring. The assumptions of the facilitated approach (as opposed to an expert approach) are in line with previous work done in the HEW project. For example, the recognition that the problem is a socially constructed, subjective entity rather than an objective reality constructed by the modeler (Franco & Montibeller, 2010), which is evident in the present study too.

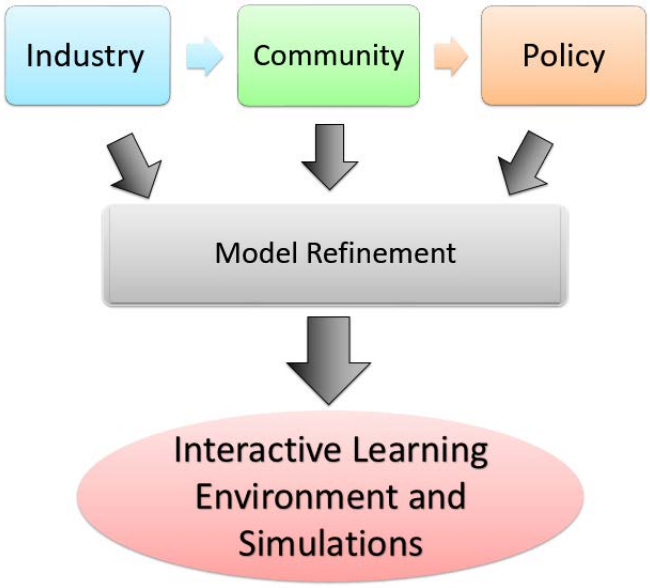


Figure 3. Overview of the engagement process undertaken. Small group workshops informed the model used to drive the simulation game.

The study was carried out in three distinct phases, shown in Figure 3. GMB workshops were held focusing on three content areas. This division was motivated by the broad problem space and highly diverse stakeholder group of the HEW project which, had so far presented challenges in arriving at a problem focus. Bringing together participants from all three levels in the same session promised to deliver very high levels of cognitive conflict. As each GMB workshop was scheduled for only 3 hours and contained an activity intended to elaborate model structure, this was deemed inadvisable based on previous knowledge of stakeholders held by UCL researchers. Therefore it was determined that dividing participants in this way would lead to more beneficial levels of cognitive conflict.

The results of these three workshops were then refined into a larger model (Eker et al., in preparation), which was then refined into the game. Notice that the game is referred to as an 'interactive simulation environment.' This was chosen out of consideration for the stakeholder's point of view that perhaps

they would find playing a 'game' too juvenile an exercise. The use of facilitated modeling techniques to produce the game was intended to boost the usefulness of the game for the wider stakeholder group.

Data collection included participant observations during the workshops as well as audio recordings. These were supplemented with questionnaires which included a measurement for learning as well as scales which were operationalized to measure impacts on fragmentation. The questionnaire also included questions to address the specific contributions of certain elements thought to contribute to group processes. Pre-test, post-test questionnaires were developed for use in the game workshop along with log-sheets, used by participants to record their thought processes. A more thorough description of the data collected and the methods used for analysis are given before the results, following a description of the design and observed outcomes of the workshops.

GMB Workshops

The primary purpose of these workshops is to apply disconfirmatory techniques to elicit structure to develop a more "adequate theory of the problem (Luna-Reyes et al., 2006, p. 304)", while supporting the overall research objectives through facilitated group process. The GMB workshop groups were smaller sub-sets of the larger HEW stakeholder group. They were invited based on their general area of expertise to represent different groups involved at multiple levels of governance or within a topic area. The model building sessions were based on best practice GMB scripts including, sticky-dots (following boundary test/variable elicitation), concept model and structure elicitation (Andersen & Richardson, 1997).

Table 2 shows the schedule or used for each small workshop. After initial introduction and greeting the stakeholders' attention was directed towards a wall in the room that contained model variables. Many of these variables were taken from the previous interview data, however, some variables were added by the modeling team while determining cause and effect relationships for the model. During the dots script, stakeholders were given three votes to distribute among the variables they deemed to be most important. This process included a large disconfirmatory element, and stakeholders were asked what variables were missing (D.L. Andersen et al., 2012). The highly ranked variables and discussion around this exercise then served as input for the structure elicitation script following the demonstration of the concept model.

Next, the concept model was gradually "unfolded" to the client group in a sequence. The sequence began by showing only a part of the structure (generally stocks and flows), then gradually introduced further variables and connections. Throughout the process, stakeholders were encouraged to ask questions in order to clarify relationships. An important aspect of the concept model is that it was not intended to be correct but to "jump start" conversation about the system from an endogenous SD point of view. The use of initially limited and even erroneous models in this way has been demonstrated to increase learning (Wijnen, Mulder, Alessi, & Bollen, 2016).

Table 3. Overview showing the activities within the small workshops.

Time	Activity	Description	Purpose
9:00	*Introduction and Agenda*	Provide an overview of the workshop, outline goals and define purpose	Inform
9:15 – 9:45	*Disconfirmatory-oriented variable identification using sticky-dots*	Divergent variable identification and convergent sticky dots	Explore model boundary, validate dynamic hypotheses
9:45-10:00	*Demonstration of conceptual model*	Divergent process, jump-start conversation	"Jump-start" participation
10:00 – 11:30	*Structure Elicitation*	Facilitate convergent structure building focusing on cause and effect and feedback	Generate ownership, encourage causal thinking
11:30-11:50	*CICC questionnaire and swing weights*	This included disconfirmatory questions about model components and *elicitation of swing weights*.	Data collection, validation
12:50-13:00	*Summary and debrief*	Give feedback on the outcomes of the day, provide stakeholders with a take home message, provide information on next steps of the project	Maintain stake-holder interest

As model structure was completed important indicators emerged from the GMB sessions and the model was elaborated. This was done via a combination of interviews and empirical data collection (Eker et al., in preparation). As multiple attributes competed for the attention of decision makers in an integrated approach to housing development, tradeoffs necessarily occur and had to be represented in the model. This need led to the use of swing weighting (described in *Data Collection*) during the policy and community themed workshops.

Swing weighting, is a technique first described by (Bodily, 1985) for eliciting the decision maker's utility for a given decision. As a part of the final questionnaire during two GMB workshops participants were directed towards an uncertain element in the model and asked to rank the group of variables affecting this element according to their relative importance. Next, the top ranked variable was given a value of 100 before being compared to the others. Participants assigned values to the remaining variables by considering the impact a "swing" from the lowest level of a given variable to its highest possible level.

Swing weighting was chosen as it is particularly effective for distinguishing the importance of a given variable in a specific decision context and prevents over-valuing of options which are of relatively little importance, compared to other options (Goodwin & Wright, 2014). The use of this technique as a script for group model building represents a novel approach when compared with existing repository of GMB scripts (Wikibooks contributors, 2016).

Gaming Workshop

The result from these small workshops were developed into a simulation model in order to drive a game, played in the final workshop. To arrive at the model, the facilitation team had elaborated and added parameter values to model structure developed from the small group workshops. This was an imprecise and difficult task, due to the number of soft variables suggested and defined by the stakeholders (Eker et al., in preparation). The resulting model had many assumptions and estimates, but its lack of validation can be an advantage in group process (Wijnen et al., 2016).

The purpose of employing a game "is to convey experiential lessons" (Lane, 1995, p. 606). In this case the experiential lessons intended to be captured were the importance of integrated approaches to meet the multiple objectives of the built environment. Both GMB participants and participants who did not contribute to the small workshops played the game, and a pre-test, post-test design was intended to compare outcomes in order to measure the changes that occur when SD simulation games are used. Theory on simulation environments emphasizes the need for transparency, simplicity and a clear description of the relationship between the model structure and its behavior. This is well established, for example (Morecroft, 1988, p. 312) asserts the following: "In order to stimulate debate, a model should be transparent so that policymakers can see their knowledge reflected in the model's assumptions. The model should also be presented in a way that dramatises its assumptions" (p. 312). This was taken into account during the design of the game, resulting in a small model containing only a few feedback mechanisms, shown in Figure 4.

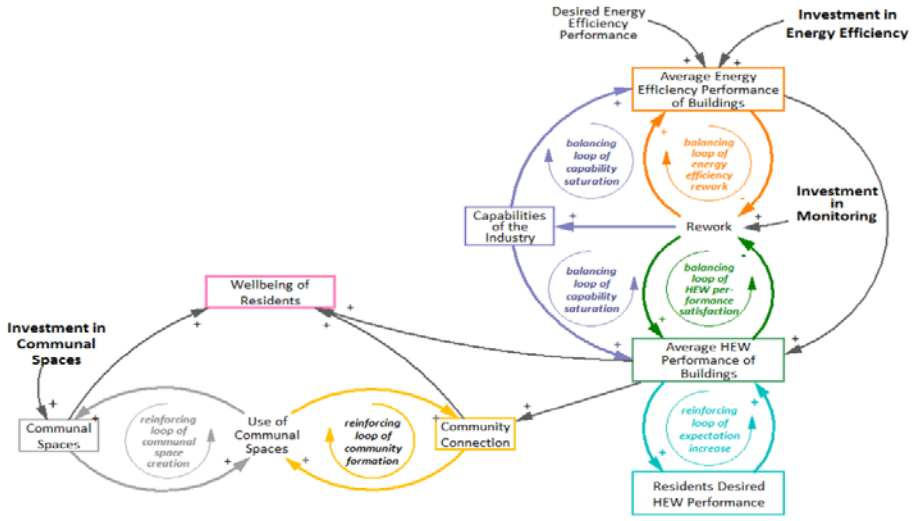


Figure 4. The CLD of the model underlying the game.

In order to motivate stakeholder engagement, as well as to provide support for learning objectives the game was played in a sequence of two different stages. This was motivated by evidence of the effectiveness of the process used by Kopainsky et al. (2015) called the prior exploration strategy. The figure below shows the two stage process applied, and is explained using the trial-driven and analysis-driven process models for game play (Van Daalen et al., 2014). In the first stage the model not visible to participants, and they were given only a brief introduction before simulating. The results of each simulation were viewed as graphs over time. The debrief (Figure 5, #5) from this stage served as a transition into stage two, where structure behind the model was presented by the facilitator. This was crucial step in the workshop as it drives at the relationship between the model structure and the simulated behavior, which has been shown to necessary to facilitate learning (Pavlov, Saeed, & Robinson, 2015). A simplified causal loop diagram was used for the demonstration to allow for better understanding of a complex system (Ghaffarzadegan, Lyneis, & Richardson, 2010). (A more in-depth description of the larger model process is provided by (Eker et al., in preparation)

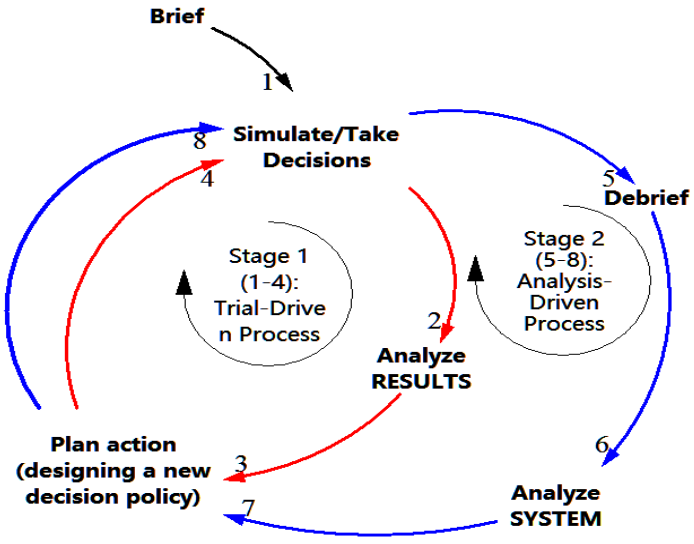


Figure 5. The stages of playing the game use trial and analysis-driven models (Van Daalen et al., 2014).

The iterative format shown in Figure 4 (i.e. multiple trial-driven simulations followed by multiple analysis-driven simulations) was chosen to increase cognitive conflict and encourage the use of the model as a boundary object for the discussion.

The use of a worksheet provided consistency across groups and follows the advice from previous SD gaming applications to emphasize debriefing in order to avoid "video game syndrome" where participants interact with the simulation environment without consideration of any learning opportunities(Andersen et al., 1990; Ford, 1996; Lane, 1995; Meadows, 1989). An emphasis on reflection by participants during the game has been deemed highly important in order for learning to occur (Andersen et al., 1990; Beall & Ford, 2011; Ford, 1996; Kopainsky et al., 2015; Lane, 1995; Meadows, 1989).

Three different investment decisions were leverage points for participant intervention. Participants' individual decisions on how to allocate funds across these decisions was intended to be used as the pre-test and post-test, to concisely measure changes in consensus. These investments related to the structure developed in the small workshops. Gameplay was guided in groups where investment decisions were required to be unanimous before simulation would be allowed. Once a decision had been reached the facilitator would input the values and run the model. The gaming log-sheets accompanied both game-play stages, encouraging decision about expected outcomes. After simulating the group answered questions found on the worksheet to stimulate reflection and discussion (*see* Appendix C).

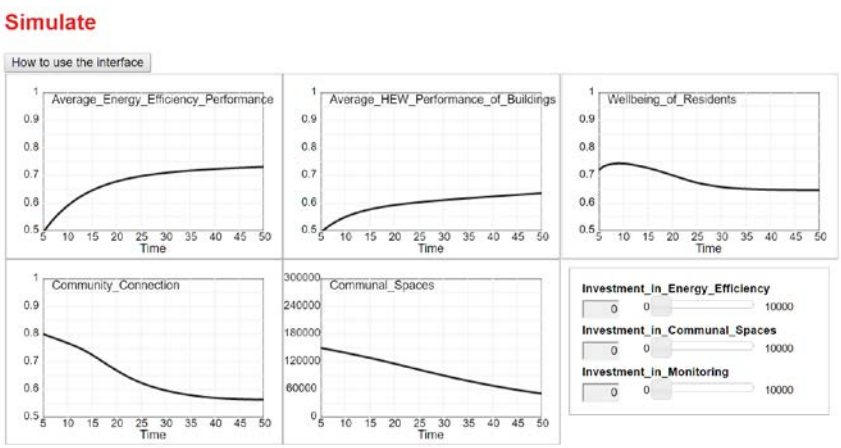


Figure 6. The main simulation interface found on the HEW-WISE website.

The model was created with Vensim® but implemented using Systo®, an online JavaScript-based simulation tool titled the "Housing, Energy and Wellbeing- Web-enabled Interactive Simulation Environment**"** or HEW-WISE (*found at:* www.systo.org/hew-wise.html). The indicators and investment decisions chosen for representation in the gaming model are shown in Figure 6. The website also includes an introduction to the project, instructions on how to use the model, a description of the causal diagrams, a frequently-asked questions section and access to the model equations. In addition to the three primary investment decisions, users are able to access and modify parameter values and graphical functions.

The results of the workshops are given below. The observational/process details of the small group workshops are given first followed by the game. Furthermore discrepancies between the intended research design and actual data collected will be explained. Next a summary of the main findings from the workshops are then provided. This includes the comparison between the two groups of participants (GMB or game) based on the CICC survey results.

Chapter 3: Summary of Observations

The previous sections outlined the study context and intended research design. In the next section the implementation of this research design is described in detail on the basis of observational data. Observational data collection was guided by the categories defined by (De Wit, Greer, & Jehn, 2012) which distinguish between relationship conflict, process conflict and task conflict. Task conflict relates to the previously mentioned cognitive conflict. It is a beneficial form of conflict which is useful in generating new ideas, leading towards group acceptance of new paradigms. Relationship or interpersonal conflict and process conflict were both considered as detrimental to group process.

Process of the Small Group Workshops

The table below shows a summary of the facilitation team's observations. The section that follows elaborates on these main points and describes key differences among the workshops.

Table 4. This table provides a summary of the observations from each small group workshop, all workshops were three hours in duration.

	Encouraging Aspects	**Concerning Aspects**
Industry n=3	• Facilitator balanced the discussion. • Cooperation among facilitators. • No stakeholder dominated and participation, engagement was high. • Participants understood concept model. • Substitute participants were useful and engaged.	• It was difficult get more than causal links out of the discussion, sometimes these lacked direction, and strength of effects discussed little. • Need for substitute participants not previously involved in HEW project. • The space available for the workshop was confining, causing some challenges in communication among facilitators.
Community n=5 (Swing weighting technique applied)	• Introduced cognitive conflict (divergent) and consensus building (convergent) aspects of GMB successfully. • New feedback loops were made and participants confirmed others. • Successful elicitation of weighted additive function via the swing weighting technique. • Disconfirmatory approach worked well for discovering weak model links.	• Some persons dominated discussions during the session. • Poor time management by facilitation team lead to overlap of exercises • Some participants seemed to struggle with putting stories into structure. • Not enough time to discuss connections to other parts of larger model.

Policy n=7 (Swing weighting technique applied)	• Convergent thinking predominated towards the end of the session. • High levels of engagement were standard. • The facilitator followed best practices during structure elicitation in the face of a challenging group dynamic.	• Pre-existing stakeholder-facilitator relationships sometimes led to confusion about who was leading the exercise. • There were several people who dominated the discussion, causing difficulty for facilitators to pace tasks. • Some participants resisted using the model to share their ideas. • The larger group size was more difficult to manage.

All three workshops followed the same basic schedule, after the project was briefly introduced and the agenda outlined, variable elicitation began. New variables were usually accompanied by a brief description of the participants reasoning. There were notable exceptions in the community and policy workshops where participant descriptions became lengthy. This was most pronounced during the policy workshop where at least half of the group tended to dominate discussions using extended story-telling as a primary method of discussion. The facilitator, when able, would relate the discussion back to the variable list but this was not always successful in identifying new variables that would be important to model. However, in most cases, the descriptions given by participants related to other variables that had been suggested by the interview results.

The end of the variable elicitation exercise was notably different among the three workshops. For the industry group, the participants expressed that they had no new variables to add to the list. In the community and policy workshops this required some pacing by the facilitator in order to keep the schedule. Likely, this effect is attributable to larger number of participants and perhaps higher levels of engagement as well. Once the variables had been recorded, participants voted on variable importance using sticky dots. This generated informal discussion among participants about the variables and issues. This was most notable during the policy workshop where one participant, in particular, was eager to begin drawing connections between variables, describing a link between learning of policy designers/analysts and competence of policy designers/analysts.

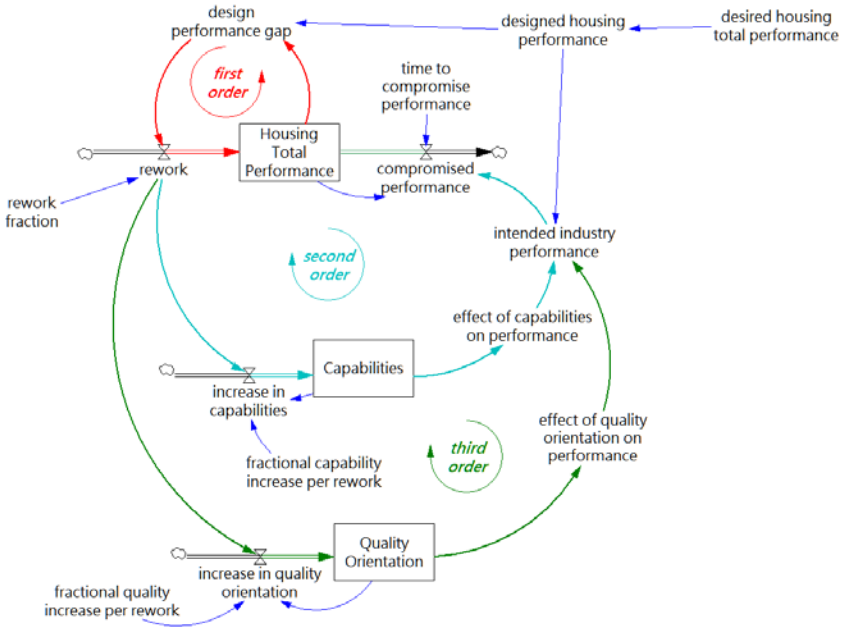


Figure 7. An example of a concept model taken from the industry workshop.

The next stage of the workshop used a concept model. In the industry and policy workshops the model was gradually revealed to the participants using the Vensim® software. In the community workshop, technical difficulties forced the facilitator to trade this gradual unveiling for a model description based on a pre-drawn stock and flow diagram appearing on a white board in the room. This took more time than was scheduled, with the facilitator checking frequently with the group for any gaps in understanding. This deviation may have been partly responsible for the observed difficulty of some participants to situate their stories as elements within the model. During the unveiling of the model and the structure elicitation exercise that followed (in all workshops) the facilitator welcomed comments and questions and emphasized that the model is a simple representation, encouraging discussion about what part of the model is wrong. Such disconfirmatory questions worked well – some causal mechanisms were validated and others were rejected. As intended, the concept models helped to kick-off discussion amongst participants.

However, creation of new structure in a coherent manner was more difficult in the community and policy sessions. In the community workshop, facilitation of building model structure focused a great deal on community connection and third spaces as well as demographic changes and gentrification issues. Participants shared rich stories, and a central variable “use of third spaces” emerged very quickly to form new feedback loops. Participants anchored to this area of the model, making it difficult to direct their attention towards other areas. This result slowed progression of model building, however it also

showed participant's emphasis and interest in the concept of third spaces and the connections between third spaces and wellbeing.

Participant engagement manifested differently during the policy workshop. Although the participants were able to follow the causal connections to understand the concept model, bringing concepts into the model in order to form new loops was more of a challenge. The facilitator directed participants back to the model structure, but not all were willing to use the model as a means to expand their ideas. Story telling was a more comfortable method for expression, either using hypotheticals or specific examples from past policy successes/failures. Again, the size of the group had a noticeable impact. A general observation was that they were able to capture loops, if the number of associated variables and connections was less extensive. Yet, they were not able to close the loops that involve a longer chain of variables. However there was one participant who stood apart, more comfortable with describing his ideas via causal structure. He offered to bring the discussion back to the model structure on his own accord and, following the workshop, sent emails to describe his hypothesis in more detail. At several points in the workshop he jumped out of his seat and came up to the board to aid in his description of loops he saw as critical.

By the end of the structure elicitation processes there was a noticeable converging of thoughts among the policy group. Participants commented on each other's narratives in general agreement, using phrases like "that's exactly the problem", instead of waiting for a pause in another's story so that they could begin their own. To close the session, the outcomes of the day were briefly reflected upon and the emphasis areas of the modeling effort so far were explained.

Process of the Game

The stakeholder schedule for the gaming workshop is shown in the table below, the full workshop schedule for facilitators can be found in Appendix B. The overall process of the day was well-paced for the most part.

Table 5. An overview schedule for the gaming workshop that was sent to participants.

Starting Time	Item
09:15	Coffee
09:45	**Welcome and the HEW Project**
10:40	**Introduction to the interactive simulation environment**
11:00	Break
11:20	**Trial-driven learning environment simulations**
13:00	Lunch
13:45	**Analysis-driven learning environment simulations**
15:15	**Exploring simulation-based scenarios**
16:00	Break
16:20	**Insights and future work**
17:30	**Closing remarks**
	Drinks and nibbles

Table 6 below, summarizes some of the overall outcomes observed by the facilitation group. Participants were allocated into four sub-groups ranging in size from three to four participants. Each of the groups groups had been assigned two facilitators. One had the facilitator role and performed tasks such as relating discussion to structure, pacing the tasks, interacting with the model to run the simulations and managing participants to limit unwanted sources of conflict. At the same time, a second facilitator was placed in an observational role. These individuals were tasked with paying attention to the interactions among participants and, as in the GMB sessions, were guided by the different types of conflict (De Wit et al., 2012). Both facilitator and observer were given written instructions defining their role and what they should watch for during the workshop. However, in the end both the facilitator and observer of each group contributed observations. This is described in full in Appendix C. A summary of the observations from the facilitators within each group is given next.

Group 1 was observed to be very cooperative despite differences in their initial positions. The observer noticed convergent behavior throughout, however it took some time for consensus to begin to emerge. The trial-driven stage of the game had the intended effect of motivating inquiry as intended. The group was very interested to perform extreme conditions tests and analyze the behavior to understand the model before testing their unanimous investment decisions. One of the participants in this group had more experience with SD and CLD diagram methods and sometimes dominated discussions. However this individual also helped trigger cause and effect thinking among the other two participants in the group. Besides this, other forms of unwanted conflict were not observed and total of three participants was manageable for the facilitation team.

Group 2 was a composed of four individuals. Discussions in this group were of a very technical nature and three of the four were very engaged. The observer noted that likely, the lack of participation by the fourth had to do with the content becoming too technical. Debate regarding the investments was generally quite intense, though respectful. However, some relationship conflict may have been involved as the observer noted comments from one participant were “a little condescending” towards others. This participant also had a tendency to dominate discussions which required management by the facilitators, and they were only partially successful at mitigating this.

Group 3 experienced some process conflict during the trial-driven phase of the game. They seemed to approach the simple model with a high degree of seriousness, and the facilitator needed to take extra time to situate the model as a tool for discussion leading to shared understanding. This group was especially critical of the model assumptions which lead to constructive discussions and cognitive conflict. In the analysis-driven portion of the game they were focused on making arguments using the causal loop diagram provided. Their criticisms of the model in this way may have served as a conduit for building shared understanding, but this had little to do with the model insights and much more to do with the discussions held during the game-play. Participants in the group also commented on the usefulness of such a modeling approach not only policy challenges but also challenges of political process.

Group 4 had help from one very engaged participant who seemed to grasp not only the task but also the process and insisted others in the group focus discussion on the available tool. All group members were

observed to be generally engaged with the game and the discussions surrounding this. However, during the trial-driven portion of the game some participants expressed reservations at making a decision without being able to see the structure of the model, representing some process conflict. They assessed that any decisions they make would be futile due to their lack of understanding regarding the model assumptions. One of the members of the group had a tendency to back up arguments by emphasizing his influence as policy decision maker.

Table 6. A summary of the encouraging and concerning aspects of the large gaming workshop.

	Encouraging Aspects	**Concerning Aspects**
Overall Workshop	• Positive discussions throughout • Good teamwork among facilitators • Simulation environment well-received • Related this workshop to other HEW work past, present and future, successfully • Interest from several stakeholders regarding use of similar technique for their problems.	• Failure to collect input data. • Too much time spent presenting previous work, this needs to be summarized more effectively.
With respect to sub-groups	• All groups tended to engage well with the model • Divergent and convergent aspects appeared within the two sessions (trial driven vs. analysis driven)	• Not all facilitators had GMB or participatory training. • Some participants found user interface to be confusing.

Following the workshop the facilitation team held a debrief meeting to share opinions resulting in the development of items in Table 5. In general the game workshop was considered a success. Facilitators and observers both reported high level of engagement from participants. Some stakeholders even approached facilitators privately, following the workshop, about using SD methodologies within their organization. They commented on the size of the model that was used for the game, noting its ease of use for non-experts. Similar conversations between facilitators and participants broached the use of SD modeling for challenges that require rapid policy response from decision-makers. Furthermore, facilitators and observers both reported high level of engagement from participants. Several noted that this engagement seemed to increase throughout the day. Unfortunately, part of the intended research design, to compare GMB participants with those who had not participated in GMB workshops, was not

realized. This was due to a failure by the facilitation team to coordinate the administration of pre-test, post-test investment surveys at the correct points in process.

Chapter 4: Results and Analysis

The previous sections reported on the observational data collected by the facilitation team during the workshops as well as the results of the swing weighting technique as used in the GMB workshops. The following section shows how this data was collected in process and describes the data sources used, including the CICC questionnaire data and audio recordings of the workshops.

Data Collection and Analysis

The CICC questionnaire played a crucial role in the study. The scales of consensus and communication provided an operational basis to measure effects of the interventions on fragmentation occurring between individual decision-makers. The research question pertaining to stakeholder learning was measured using the scale for insight, as indicated by (Vennix & Rouwette, 2000; Vennix, Scheper, & Willems, 1993).

*Table 7. An overview of what data was collected and when. *Pre-test questionnaires were distributed but not filled out by all participants.*

Data Collected	Point in Process
Consensus, Insight and Commitment to Conclusions (CICC) Questionnaire	After each small workshop, and after gaming workshop.
Investment decisions	Before* and after gaming workshop
Gaming log-sheets (includes group investment decisions)	During the game
Observational Data	During each workshop
Audio Data	During each workshop
Swing Weights	During community & policy workshops

In addition to the items pertaining to these four scales, participants were also asked to compare the workshop (either GMB or game) to a normal meeting. Questions about specific elements were also addressed. A model validity section was added to the CICC questionnaire, and was modified for each workshop in order to address specific variables, sourced from previous interview data (Eker & Zimmermann, 2016). In two of the GMB workshops, tables to facilitate swing weighting were also included, these are found in Appendix C. The questionnaire from the gaming session also included additional questions about participants' policy priorities, their consideration of the method as useful and a field for additional comments.

Data collection was carefully considered with the aim to gather as much information as possible without exhausting the participants, many of which held demanding and high-level positions relating to the built environment. As previously described the pre-test, post-test design was intended – however it was not realized. Gaming log-sheets were designed to guide facilitators of the game, ensuring that participants

would carefully consider each investment and their intentions before running the simulation and analyzing the result. The disconfirmatory approach, based on D.L. Andersen et al. (2012) was used throughout the project. This strategy sought to overcome the deference effect, or respondents telling the interviewer what they think they want to hear, by explicitly asking participants to look for errors or unimportant variables in the model. This was applied most directly in the small group workshop settings and in the "model validity" portion of the CICC questionnaire which included a bank of items regarding usefulness of different model elements. The recurring theme of limited stakeholder access confounded further use of this strategy in individual interviews. For consistency the facilitation team contributed to a debrief session after each workshop where the encouraging and concerning aspects were discussed, as well as how to improve for the next session.

Table 8. An overview of the more detailed coding manual developed for the analysis of transcribed audio data is situated within the CICC questionnaire. Additionally, frameworks to assess boundary object use and the ability of the process to support consideration of soft variables, was also developed and applied.

Research Objective	Method and Data Source	Positive Codes	Negative Codes
Learning and insight	CICC and audio data	• Inquisitive statements regarding counterintuitive results • Statements of surprise • Describing feedback • Explicit statements of understanding	• Rejecting counterintuitive results without inquiry • Describing relations as linear • Explicit statements of not understanding
Boundary object in game	Stage 4 of (Black & Andersen, 2012), audio data, CICC - commitment	• Modification of visual element • Participants make reference to visual element to construct an argument • Statements about tool usefulness, in present and in future applications. • Expressed interest to share results or ideas from workshop with others.	• Participants avoid discussing the visual element • Causal descriptions that don't relate to the visual object • Expressed frustration with visual element (e.g. inability to change as desired) • Expressing doubts about model uses in present or future. • Expressed difficulty using tool

Boundary objects in GMB workshops	Stage 1-3 of (Black & Andersen, 2012), audio data	• Modification of visual element • Participants make reference to visual element to construct an argument • Statements about visual element's usefulness. • Discussion of interrelationships and dependencies • Addition of new variables	• No additional variables proposed • Viewing variables as isolated or not connected • Rejection of terms used by others. • Participants avoid discussing the visual element • Arguments that don't relate to the visual object • Expressed frustration with visual element (e.g. inability to change as desired)
Fragmentation	CICC – consensus & communication	• Converging of opinions, agreement, and accommodation of other points of view. • Common language use • Participants share speaking turns	• Disagreement, unwillingness to accommodate other points of view • Use of different vocabulary • Participants give multiple lengthy descriptions
Consideration of multiple objectives	No particular method, audio data	• Positive statements regarding soft variables • Agreeing on definitions for soft variables • Expressed desired to have higher performance on multiple indicators and soft variables	• Negative statements regarding soft variables. • Disagreement about definitions for soft variables • Concern with one indicator only

These measures were further augmented by audio recordings that were taken during the sessions. These recordings were then transcribed and analyzed. Table 7 shows a summary of the coding manual used. It was developed via a combination of inductive and deductive process (Franco & Rouwette, 2011). Deductive coding drew on descriptions of consensus and communication, given by Rouwette (2011), and were used to relate the qualitative coding to the quantitative CICC questionnaire and provide a means of operationalizing the identification of reduced fragmentation among stakeholders. An item measuring communication: "The modelling process aided in the understanding of the opinions of the other participants." The emphasis this item placed on understanding and consideration of other's points of view directly addresses the fragmentation occurring between individual decision-makers. As do items on the scale of consensus, for example: "Our opinions are closer due to the modelling process." Learning, as described by (Vennix & Rouwette, 2000, p. 200) relates to the scale of insight, an example

item from the questionnaire: “The modeling process has given me more insight into the feedback processes that play a role in the problem.” As shown in the table, descriptions that describe feedback are positive indications of learning. Negative indications are coded when relationships among variables are described in a linear fashion. Therefore the coding places emphasis on the relationship between each quantitative scale and the qualitative coding analysis.

The coding was completed by the author and benefits in consistency from in that regard. However, it is possible that in the process of coding the large volume of raw data ‘coder’s decay and drift’ could have influenced the result. This refers to the understanding of the coder changing as they progress through the data (Folger, Hewes, & Poole, 1984). This can be mitigated in a few ways to improve the reliability, including iterative re-training and frequent referral back to a written manual. The latter of the three was used in this case. However, due to time constraints, only one iteration was conducted with a second coder who was unfamiliar with the study. Only a small portion of the transcripts were checked in this way (5 units of 120 total units, see: *Appendix* B). This was used to calculate a *Cohen’s kappa* value of 0.21 (Cohen, 1960). This value is quite low, as a kappa of 0.70 is generally considered satisfactory. Comments from the second coder were used to revise the code manual, however future revision and re-training is needed to improve the reliability of the result.

The framework given by Black & Andersen (2012), shown in Figure 2 of the literature review, was operationalized to provide a basis for assessing the function of the visual elements used as boundary objects in each of the two workshops. The GMB workshops supported stages 1, 2 & 3, where two visual elements were combined. First the disconfirmatory dots exercise, and second, a concept model plus elicited structure. Satisfying the requirements of a boundary object, the disconfirmatory dots exercise was a visual representation of variables and facilitators encouraged all group members to add new variables. The discussion of emerging variables along with the sticky-dots portion of the exercise encouraged thinking about dependencies, as participants had to consider the relative importance of the related variables. This was expected to correspond closely with the *generating ideas* process shown in the figure. The concept model was the starting point for structure elicitation which, taken together, also fulfills the requirements of a boundary object. Participants could add variables and the exercise emphasized the elaboration of cause and effect thinking within the model structure, providing a close match for *identifying dependencies* in ideas. Discussions surrounding each of these elements supported the *transformation of ideas*. This allowed for user input on the model and focused on the links between the variables. The resulting causal structure represented transformed ideas in the form of new feedback loops.

Stage 4 should then be supported by the game, used in the final workshop. The game fulfills all necessary characteristics of a boundary object: It is a visual representation of the relationships among policy criteria and it is transformable by all participants via decisions which influence parameter values resulting in different simulation results. The design of the workshop which forced participants to prioritize funding allocations directly corresponds with the *selecting ideas* phase. Success or failure of the game to serve in the role of a boundary object, and therefore to support the development of stage 4, provides the measure of successful integration of GMB and games in this study.

To underpin the analysis of the game this was combined with the CICC questionnaire scale of commitment which is defined in terms of either a decision or other desired result from the intervention (Vennix et al., 1993). An example item from the questionnaire is: "I will uphold the conclusions/findings of these meetings in front of other members of my organisation." Participants agreeing with this statement are showing their approval of their group's findings. Therefore this aspect of the game as a boundary object was considered in light of the purpose of this intervention, which was not intended to reach policy decisions, but rather encourage integrated thinking and coordination among stakeholders in the built environment.

Results

The CICC survey data were compared between the small and large workshops following the examples of Eskinasi & Rouwette (2004) and Vennix & Rouwette (2000). Questionnaire data was collected from 13 stakeholders who participated in the small group workshops. One participant attended the policy and community workshops and is therefore represented twice in the analysis (n=14 in Tables 9, 10 and 11).

The table below shows the results including min, max, mean and standard deviation or the four scales. A 5-point Likert scale was applied ranging from disagree (1) to agree (5). For each scale a two way t-test was used to compare the means of the survey result to a neutral score (neither agree nor disagree, 3). The mean for the results was found to be significant for all scales in both workshop types (*t-test 2-tailed significance <.000)* and, results *between* the two groups were also found to differ significantly (*t-test 2-tailed significance, independent samples < 0.000*). Therefore, both meeting types had positive effects on communication, insight, commitment and consensus. For the GMB workshops insight and commitment were significantly higher. In the game workshop consensus and communication were greater.

*Table 9. Final results summarized, the results of the workshops are significant and positive, however coefficients are above the threshold value (0.60), after applying the Spearman-Brown** prediction formula. The gaming workshop had a significantly greater positive effect* on insight and commitment than the GMB workshops, which performed better on consensus and communication scales.*

	# items	α**	n	Min	Max	Mean	Std. dev.
Game Workshop							
Insight	4	0.74	9	3.25	5	4.31*	0.65
Consensus	3	0.70	9	2.34	5	3.85	1.01
Commitment	3	0.91	9	2.67	5	3.79*	0.86
Communication	3	0.75	8	2.66	5	3.79	0.86
GMB Workshops							
Insight	4	0.62	14	3.25	5	4.19	0.77
Consensus	3	0.88	14	2.67	5	4.12*	0.85
Commitment	3	0.87	14	2.34	5	3.62	0.86
Communication	3	0.87	14	2.34	5	3.98*	0.9

As noted in Table 8, not all tests of reliability were above the threshold value of .80, after a conversion using the Spearman-Brown prediction formula. This formula was used to calculate the reliability of each scale at if it had been extended to 10 total items. The survey has been used with a larger respondent group, where a reliability of .82 was found (Vennix & Rouwette, 2000).

In addition to the four scales, questions asked for a comparison between the workshop they experienced and normal meetings. The results shown in Table 5 show that both workshop types were better on all dimensions over the stakeholder's idea of a standard meeting. However, no significant difference was found between the GMB and game workshops.

Table 10. A comparison of normal meetings to the workshops found no significant difference between those who attended the small GMB workshops and the larger gaming workshop. Scored on a scale of -5 to 5.

	Game			Small Workshops			P-value
	Mean	*Std. dev.*	*n*	*Mean*	*Std. dev.*	*n*	
more insight	4.44	0.88	9	4.36	2.98	14	0.80
faster insight	4.11	0.93	9	3.71	3.13	14	0.36
better communication	4.33	1.12	9	4.50	2.83	14	0.69
faster alignment of mental models	3.89	1.17	9	4.21	3.37	14	0.76
better alignment of mental models	4.11	0.93	9	4.36	3.72	14	0.75
faster commitment	3.67	0.87	9	3.79	4.13	14	0.32
more commitment	4.00	1.00	9	3.79	4.04	14	0.61

The results shown in the table below are based on specific elements of GMB. The means for the gaming workshop were higher than for the GMB workshop leading to two significant differences.

*Table 11. A comparison of the contribution of specific GMB elements toward the workshop success. Scored on a scale from -5 to 5, *denotes a significant difference among the two workshops.*

	Game			GMB Workshops			P-value
	Mean	*Std. dev.*	*n*	*Mean*	*Std. dev.*	*n*	
projection of diagrams	3.88	1.46	8	2.71	2.61	14	0.20
presence of a group facilitator	4.22	0.67	9	1.71	2.79	14	*0.00
opportunity for discussion	4.44	0.88	9	4.14	1.10	14	0.48
use of causal loop diagram	4.00	1.66	9	3.29	1.54	14	0.32
computer model simulations	4.33	1.00	9	1.09	1.04	11	*0.00

In addition to the CICC data audio recordings were transcribed for both workshop types and coded beginning at the first indication of use of visual elements in the process and ending at the close of scheduled workshop activities. Only the second and third GMB workshops were recorded following the realization by the facilitation team that this would be a useful supplement for model revision. Due to time restrictions only three groups from the game workshop were transcribed. This was chosen on the basis of the facilitator skill level, in order to better investigate the effect of this variable on outcomes. Group 1 had a highly-skilled facilitator who had been involved in the HEW project for more than a year. The facilitator for group 2 was representative of an intermediate level of skill, having gained some facilitation experience through involvement in the previous GMB workshops. The group 2 facilitator was relatively new to the HEW project, but had been involved throughout the process and was the main developer of the model. Finally, the group 3 facilitator represented a low-skilled facilitator who was new to the project. This facilitator had many years of experience applying SD in an expert fashion but lacked GMB training and experience and was not involved in the HEW project at the time of this study. The group 4 facilitator shared similar characteristics with that of group 3, and was not transcribed. Time

restrictions also informed the unit of analysis with the bottom half of every page representing one unit. Therefore the analysis includes half of the total transcribed data. The results are shown in Table 11 and were analyzed using a two-tailed binomial non-parametric test with a proportion of .50. Results were reported on the basis of a .01 significance level, due to the small sample.

Table 12. This table shows the results of the audio data analysis for each individual workshop, including the proportion positive (in parenthesis) and significant positive outcomes.*

	GMB Community	GMB Policy	Group 1 Game	Group 2 Game	Group 3 Game
Transcript Length	1:58	2:40	1:34	1:42	01:33
Number of Participants	5	7	3	4	4
Multiple Objectives Positive:	12 (1)*	10 (.84)	15 (1)*	36 (.95)*	19 (1)*
Multiple Objectives Total:	12	12	15	38	19
Learning (Insight) Positive:	4 (1)	8 (1)*	21 (1)*	37 (.95)*	10 (.63)
Learning Total:	4	8	21	39	16
Fragmentation Positive:	20 (.87)*	36 (.76)*	22 (.88)*	28 (.74)*	17 (.85)*
Fragmentation Total:	23	47	25	38	20
Boundary Object Positive:	20 (.77)*	41 (.64)*	21 (.63)	26 (.55)	29 (.62)
Boundary Object Total:	26	64	33	47	47

Overall, the results are positive and significant. Indeed, this result was seen for all workshops on the basis of fragmentation, corresponding to the significant results found for the consensus and communication scales of the CICC. The analysis shows one categorical difference regarding the effect of the boundary object between the GMB and game workshops. In all of the game workshop groups no significant positive contribution was found. The ability of the game to function as a boundary object was the primary measurement for successful integration of the two methods. Therefore, this finding does not support the use of the two methods in combination. To further investigate this difference the trial and analysis-driven portions of the game were analyzed using the same approach and results are shown in Table 13.

Table 13. Table showing results of the trial and analysis-driven portions of the game workshop, including the proportion positive (in parenthesis) and significant positive outcomes.*

	Trial	Analysis
Multiple Objectives Positive	44 (.95)*	26 (1)*
Multiple Objectives Total	47	26
Learning Positive	36 (.97)*	32 (.82)*
Learning Total	37	39
Fragmentation Positive	39 (.83)*	28 (.78)*
Fragmentation Total	47	36
Boundary Object Positive	36 (.65)	44 (.63)*
Boundary Object Total	57	70

Both stages exhibit significant positive effects on the basis of multiple objective consideration, learning and fragmentation. However, the boundary object was found to have a significant positive effect only in the analysis-driven stage. Though the proportion of positive responses (63% for analysis-driven) was lower than found in the trial-driven stage (65%).

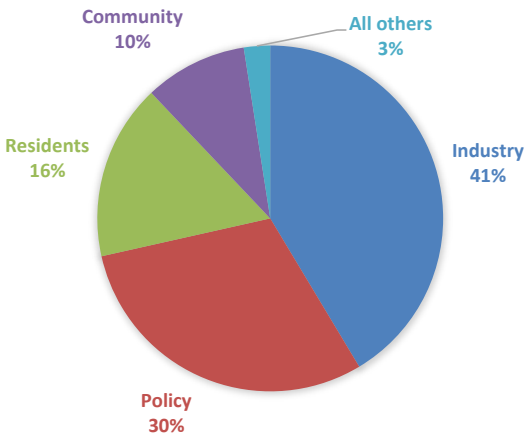


Figure 8. Weighting of inputs that increase implemented HEW performance.

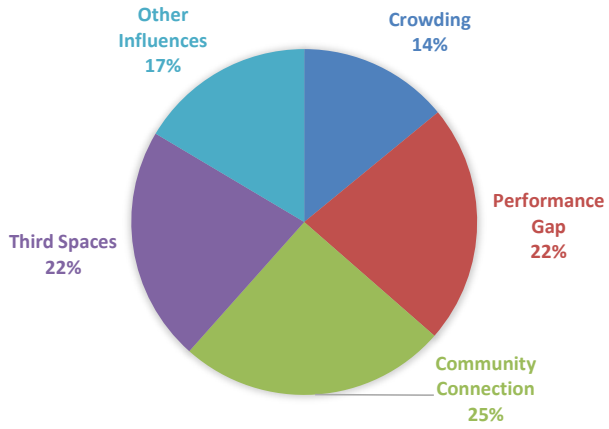


Figure 9. Weighting of inputs that increase perceived physical, mental and emotional wellbeing.

Finally, the results of the swing weighting exercise are shown in Figures 8 and 9, these were used as the basis for formulating equations for the highly aggregated variables used in the simple model underlying the game. The production of these values within the GMB workshops shows that weighting techniques can be used to elicit stakeholder values in GMB workshops. Questions of usefulness of this approach depend on interpretation of the qualitative data which will be addressed in the next section.

Chapter 5: Conclusion

The applied objective of this study centered on improving stakeholder capacity for integrated decision making in the built environment of London, U.K., in order to address the problem of fragmentation and improve stakeholders' consideration of the multiple objectives of housing. It also aimed to contribute to current theoretical understanding of the relative contributions that group model-building and system dynamics based games have on group processes. In order to do so, these two methods were integrated, and compared, with regards to their ability to improve participant learning, communication and consensus. In addition, a comparison was made on the basis of how well visual elements used for each intervention type functioned as boundary objects that support group process. Finally, the swing-weighting technique was used to elicit stakeholder perceptions of the relative influences of certain model parameters on others. The aim of which, was to simply demonstrate whether or not this technique could be constructively applied within GMB workshops.

In this chapter, results for each of these objectives are interpreted, followed by a discussion of the limitations and implications for future work. In the sixth, and final chapter, the author's reflections are given, which may help situate some of the limitations within the context of the research process.

Interpretation of the Results

The results in Table 11 provide general support for the CICC outcomes and help to elucidate where differences existed. Most of the coded measures corresponding to the scales of the CICC questionnaire were also found to have significant, positive effects on the process. For example, the fragmentation measure which was developed in connection to the consensus and communication scales of the questionnaire is also significantly positive across all groups.

The significantly higher result for the game regarding insight is not directly supported the audio data. However, it is possible to explain the discrepancies in light of the audio data. The insignificant result on this dimension for group 3 may have been masked in the aggregated CICC questionnaire results as both groups 1 & 2 show a high frequency and proportion of positive responses. This discrepancy may be explained by the skill level of the facilitator, who was the most inexperienced of the three. Examples from the transcribed data support this notion as well, since the facilitator of that group directly contributed to content, presenting his own ideas about model assumptions. This goes directly against the recommended GMB practice of facilitator neutrality (Vennix, 1996). For example, when one participant disagreed with his idea he became defensive and argumentative. The participant, who also became defensive, eventually stated, "I'm *just* questioning relationships", before raising his voice and providing a lengthy counterexample. Exchanges such as these were frequent and, though based around the model, they likely inhibited learning.

Additionally, the insignificant outcome for the learning measure in the GMB community workshop also supports the difference found in the CICC questionnaire as does the observational data. For example, participants appeared to struggle with putting ideas into the causal structure and in describing wider system feedback effects during the GMB community workshop. Also, relative to the industry and policy workshops, where the facilitator ran out of room to expand the diagram, the structure that resulted

from the community workshop was much smaller and less elaborate. Despite the significant result for the GMB policy workshop, a low number of frequencies was found for both GMB workshops, relative to the frequency found in the game workshop.

The higher score for insight can also be interpreted as a result of the 2-stage design, or *prior exploration* strategy, performing as intended. This evidence reinforces the application of such a strategy to improve learning outcomes previously shown by (Kopainsky et al., 2015). Therefore, with regards to insight & learning, the results seem to point to the game workshop as better at supporting learning outcomes.

The results show that the two stage design of playing the game also has implications for its effective use as a boundary object. As described previously, a key element of boundary objects is that they portray "dependencies and relationships among participants' objectives, expertise decisions and actions (Black, 2013, p. 200)." This approach, which kept the model hidden from sight prevented participants from directly viewing the interrelationships between variables during the trial-driven phase, this is a likely reason for the difference in effectiveness. Statements during the trial driven phase support this as well. For example, a participant in group 2 said, "I suppose the thing that troubles me slightly is that we are experimenting with an unknown structure here, the structure of the model." The web-enabled interface was also a contributing factor as it was cumbersome for facilitators to use. The audio data for group 3 show this, as the game the facilitator failed to save several of the simulation runs, and had to take time out of group interactions to re-run the model. In addition, there were several instances in all groups where participants could not distinguish the runs from one another without considerable scrutiny. Participants may have vented their frustrations at the game as a visual object however, the game did indeed support higher levels of learning. Therefore this result can also be interpreted as a failure of the coding to adequately capture the boundary object function of the game model during the trial driven stage. This juxtaposition, between stakeholder frustration with the model and positive group outcomes is demonstrated by one stakeholder's feedback from the final questionnaire that read; "The quantitative model has limited use, it is the discussion around assumptions and relationships which matters."

With respect to encouraging participants' consideration of the multiple objectives of the built environment, all but the GMB policy workshop were positive and significant. This discrepancy is could be due to participants anchoring their ideas to the concept model structure, where community connection was included only as an external influence. Once again sample size may be to blame, since the indications from the transcribed data show strong evidence of broader thinking about built environment performance. For example one participant describing a "fundamental problem" said, " energy efficiency policy may be trying to tackle poor efficiency and fuel poverty and wider issues like climate change[…] but they're not thinking of the social and community things[.]" However, this result may also be attributed to the more homogenous mix of participants resulting from the decision to group the workshops around a general area of expertise.

The significantly higher contribution of the group facilitator for the game workshop may be explained by a number of factors. First, the overall content knowledge of the facilitators and observers was higher in the game workshop, since the workshop involved facilitators and observers from within UCL's built environment faculty with many years to decades of experience. This is in stark contrast to the facilitator

group for the GMB workshops, who's most experienced member had no more than one and a half years of experience in this domain. Second, the fact that the final workshop clearly defined a facilitator and observer who remained with the same group for all game-play portions of the workshop. This was not the case in the GMB workshops where the primary facilitator changed between exercises, which may have made the facilitator role less obvious to the participants. Another contributing factor was the pre-existing relationships between a member of the facilitation team and the stakeholder group which observed challenges for facilitation, noted in Table 3 under 'concerning aspects'. The transcript supports this observation, where participants directly address this member of the team, rather than the facilitator who is guiding the exercise of structure elicitation. Additionally, the higher score for the gaming workshop with regards to computer model simulations makes sense given that there was no simulation element in the GMB workshops. On the other hand, this was the primary focus during the game workshop.

The use of swing weighting as a group model building script was perceived a useful by the facilitation team. The technique was piloted in the community workshop successfully, which lead to its re-use in the policy workshop. The resulting weights are shown in the previous section, and an example of how this was implemented on the questionnaire is given in Appendix C. Some stakeholders were able to perform the task after being given instructions only once. However, others required further explanation and assistance in order to grasp the concept. In the latter case, metaphors of other simple decision problems (e.g. consumer purchases) were helpful in getting the point across. Similar use of metaphors has been demonstrated previously. For example, the best practice concept model script where a bathtub serves as a metaphor to introduce clients to stocks and flows (Luna-Reyes et al., 2006). This should be included for further application and development of the swing weighting script. The ability of the swing-weighting technique to prevent stakeholders from over-emphasizing relatively unimportant options (Goodwin & Wright, 2014) can be a boon for modeling in subjective problem spaces.

In this applied case study it appears that the both workshop types contributed to solving the issue of fragmentation among the participating stakeholders by fostering learning, communication and consensus among the stakeholder group. Group model-building was used in combination with a simulation game to encourage the involved stakeholders to address multiple objectives of the built environment, with a focus on social and individual wellbeing indicators. This objective was broadly supported by observations, specific stakeholder feedback and coding of transcribed sessions.

Perhaps most interesting is the contribution this study makes on the basis of visual elements used as boundary objects. It demonstrated the way in which a theoretical framework (Black & Andersen, 2012) can be operationalized to support analysis of the use of boundary objects in order to assess the integration of GMB with SD-based games. This assessment method provided supporting evidence of the positive role played by visual objects used in some GMB scripts (Black, 2013; Richardson, 2013). Indeed, this study's approach to analyzing the success of visual objects to function as boundary objects may be a useful addition to standard reporting guidelines that have been suggested for GMB (Rouwette et al., 2002) and can provide a means for formally investigating whether the visual elements some GMB scripts are more effective than others. The effectiveness of the game at achieving positive group outcomes on the same dimensions and based on same questionnaire as the GMB workshops, despite the lack of

support of its function as a boundary object in the audio data, suggests that the definition of boundary objects based on current theoretical understanding could be in need of further revision.

Another interpretation is that the impact of the game is achieved in a different manner than a boundary object. Take for example a recent study by Martin et al. (2015) describing the development of social stress that occurs when a person comes in contact with unfamiliar people. Exposure to unfamiliar people results in a 'fight or flight' response, which in turn, blocks the neurological pathways that generate an empathetic response. They demonstrated that playing games can be a way to reduce social stress and therefore increase empathy. To test this, two experimental groups were subjected to a painful stimulus, 1) in the presence of a stranger and 2) in the presence of a friend. Those experiencing the stimulus with a friend reported significantly higher levels of pain than did those sharing the experience with a stranger. *However, this difference disappeared when strangers engaged in only 15 minutes of playing a game together*. Playing the game alone, on the other hand caused no change in a subjects experience of pain around a stranger. If collaborating in game play can cause an individual to 'share' a stranger's pain burden, by reducing social stress and therefore increasing empathy, perhaps a game can also help motivate decision-makers to better share each other's perspectives regarding policy priorities.

Furthermore, the study adds to theoretical knowledge regarding how SD-games can be used with groups (Bassi et al., 2015; Eskinasi & Rouwette, 2004; Ruud & Baakken, 2003) to facilitate individual learning but expands this to other important outcomes of group processes, namely the ability to generate shared understanding or consensus. Understanding the process elements that contribute to positive participant interaction with games can help practitioners design more effective methods of game-play. Finally, the elicitation of weights during GMB is still in a beginning phase but could be used to elucidate the relative effects of different variables of GMB exercises in a more objective fashion, resulting in a model which more accurately reflects stakeholder values.

The conclusions based upon these results must be interpreted in a precautionary manner. The small sample size and unforeseen barriers encountered during the research process resulted in time constraints that prevented a more thorough analysis of the data. (This is elaborated in more detail in the *Reflections* chapter that follows.) Still, these outcomes can help guide future stakeholder engagements and research strategies.

Discussion and Limitations

The applied nature of this research resulted in several limitations. The many high-level stakeholders involved made it difficult to secure participation and anticipate drop-outs, led to the small sample size. The final gaming workshop in particular suffered from many last-minute drop-outs from the stakeholder group and this may have been due to the political chaos that ensued following the EU referendum held in the U.K ("EU referendum - GOV.UK," 2016), which took place in the week leading up to the game workshop. Furthermore, as this workshop involved a larger number of participants, more administrative and coordinating tasks were necessary leading up to the day of the workshop. This was likely a factor in the previously mentioned failure to collect pre and post-test surveys as was intended in the research design. Application of the more rigorous pre-test, post-test design should be used in future comparisons to overcome the known limitations of this studies' strictly post-test survey results (Rouwette et al., 2002).

The boundary object results pertaining to game use should also motivate further inquiry. Since the outcomes for consensus and learning were still positive, despite the poor performance on this metric overall. Therefore a possible tension between the prior exploration strategy to promote learning in games and the use of boundary objects to accomplish similar goals in GMB. The findings of this study do make the case that practitioners should consider carefully when combining the two. Future research could investigate the integration of GMB and games without a distinct trial-driven phase. Perhaps the GMB element can serve as the prior-exploration phase offered by (Kopainsky et al., 2015).

Building on this, future research should pursue a more direct comparison of game play between participants with no involvement in GMB process, and perhaps include a scale to measure the levels of empathy or social stress experienced by individuals (Martin et al., 2015). The applied nature of this study was a limiting factor in accomplishing that goal. Therefore research comparing the two approaches could benefit from further experimentation in controlled settings.

Demonstration of the online tool provides an avenue for future dissemination of socially important issues. This was not fully leveraged for its potential contributions to this research. For example, a simple visitor count feature and feedback field would have provided an avenue to gain more useful data. In future applications, research should consider the potential for augmentation that such a platform provides.

During the course of this thesis research the author was appointed the title of 'Honorary Research Associate' and employed by UCL as an external consultant to support the creation of the online web-based tool as well as train facilitators and develop detailed action plan for the game workshop. Funding was provided by UCL Innovation and Enterprise, an entity that "enables knowledge to be exchanged between UCL researchers and the potential users of that research, with a view to maximising its economic and societal benefits."("UCL Innovation and Enterprise," 2016) Conducting research in light of this stated aim did not result in any conflicts of interest. The author strived for objectivity and throughout the project. All stakeholders signed consent forms which guaranteed their confidentiality in the process and outcomes of this work. Permission was obtained for all audio recordings. In some cases, administered surveys intended to track the changes in individual participants asked for participant

names. However, they were also given the option to omit this information entirely or use a self-determined identifier if they wished to remain anonymous. In some cases, participants were reimbursed for travel expenses to the workshops but they received no form of payment for their participation in the workshops.

Chapter 6: Reflections

In this brief chapter I will describe the overall setting in which this work was carried out in. As previously mentioned, this is an important component due to the multiple ways in which it influenced the production of this thesis.

An important component of the European Master Program in System Dynamics (EMSD) relates to the international emphasis of the curriculum. On the program website, objectives of the program are described:

> *"Our goal is to teach you everything you need for starting an international career in strategic modeling with System Dynamics* ("Objectives | European Master in System Dynamics," 2014)*."*

This international emphasis motivated my decision to engage in a project at the University College London (UCL). In addition I was driven by the opportunity to integrate with a project which had an established stakeholder group. This was the first time any student from the EMSD had engaged in such a partnership with UCL which held implications for the administrative difficulties. The navigation of these challenges was a useful learning experience but also had detrimental effects on portions of the research. Furthermore, the referendum in which the U.K. voted to exit the European Union which took place on the 23rd of June also may be to blame for the reduced attendance that affected the outcomes of the final workshop.

A first challenge was arranging housing in the high priced London market. This was a significant distraction from my studies and it took two weeks before I managed to arrange suitable accommodation. To offset the high costs of living I had collaborated with my adviser to compete for a funding bid which would be used to supplement the duration of my stay. The bid was successful, however, the administrative channels by which I would receive this funding was uncertain. Again, I took time away from research to have meetings with HR staff and after 2 months it remained uncertain as to when, or if, I would receive the funds. This uncertainty, coupled with the high costs of living made it financially infeasible for me to remain in London. Eventually, it was determined that I could be employed as a U.S. citizen however, this came too late and I was forced to give notice to my landlord. I then relocated to Lisbon where I would benefit from access to university facilities and a lower cost of living. In this way I was able to recover a portion of the costs incurred during my 3 month stay in the U.K.

Administrative difficulties resurfaced upon my return to London as well. As I was now employed by UCL as a U.S. citizen I was not allowed to do any paid work in the U.K. without violating visa laws. I was in the U.K. on a tourist visa for both the three small workshops and now for the subsequent return for the gaming workshop. When I arrived for my second visit, my primary reason was to conduct research, therefore I was unknowingly violating a limitation on visitor visas. I had been completely transparent with the HR department regarding my activities and provided any required documentation, including passport copies and academic records. However, it wasn't until I had returned that HR made me aware of these requirements. Emails were sent to myself and my UCL adviser stipulating that I was not allowed to remain on UCL's campus for more than a brief meeting with my adviser. I expressed my concern

about this to my adviser, who addressed HR. Unfortunately, this did not resolve the issues and during a meeting with the project team I was asked by an HR representative to leave the premises. When my adviser questioned this action, I was threatened with removal by security. I packed my things and was escorted out of the building.

This was disappointing for two reasons. First, as mentioned previously, I had been clear about my intentions to study and live with administrative and academic departments from EMSD and UCL from the very beginning. No one had raised any of these issue in the months leading up to my departure for the U.K. Their failure to recognize the legal impossibility of my involvement with the project had placed me in a vulnerable situation. However, I also bear some of the responsibility as I too could have further investigated the administrative requirements more thoroughly.

Second, now that I could not be on UCL's campus I would not allowed to attend the final workshop. As I was the main designer of the intended research design I had an important role to play there. I was tasked with training the facilitators, and had defined the roles. My presence at the workshop was then needed to ensure that these instructions were clear and that they were followed. This was also a critical data collection opportunity. Being absent, I could not collect observational data first-hand, nor could I ensure that the requisite questionnaires were distributed correctly.

In addition to my exit of UCL's premises, there was another exit of a political nature which had consequences for final workshop, which was scheduled just seven days after the historic 2016 "Brexit" ("EU referendum - GOV.UK," 2016). I had arrived in London one day before the vote took place in order to allow time to make final preparations and coordinate activities with the rest of the team. After the vote, London was in a state of unrest and uncertainty. The effects on British society were far reaching, which likely occupied the already scarce availabilities of many of our stakeholders. More than 30 participants were invited to participate, and our team anticipated a turnout of around 20 individuals. Instead, the final workshop was attended by 15 participants and only nine of these remained for the full workshop (see: *Process of the Game*).

These cumulative experiences, good and bad, all provided opportunities for learning. I will tread more cautiously in the future when establishing new research and project partnerships across borders. I was also reminded of the important role that implementation and coordination has alongside the research design itself. Furthermore, I now fully appreciate the very real impact that the broader political landscape can have on ongoing research. I came away with lessons about working internationally that I will apply throughout my career. In this respect, despite the challenges, these experiences exemplify the fulfillment of the objectives of the EMSD.

References

Alessi, S., & Kopainsky, B. (2015). System Dynamics and Simulation/Gaming Overview. *Simulation & Gaming*, *46*(3–4), 223–229. http://doi.org/10.1177/1046878115596390

Andersen, D., Chung, I. J., Richardson, G. P., & Stewart, T. R. (1990). Issues in designing interactive games based on system dynamics models (Vol. 1, pp. 31–45). Presented at the Proceedings of the 1990 International System Dynamics Conference, Chestnut Hill.

Andersen, D. L., Luna-Reyes, L. F., Diker, V. G., Black, L., Rich, E., & Andersen, D. (2012). The disconfirmatory interview as a strategy for the assessment of system dynamics models: *System Dynamics Review*, *28*(3), 255–275. http://doi.org/10.1002/sdr.1479

Andersen, D., & Richardson, G. (1997). Scripts for group model building. *System Dynamics Review*, *13*(2), 107–129.

Bassi, A. M., Rego, F. D., Harrisson, J., & Lombardi, N. (2015). WATERSTORY ILE A Systemic Approach to Solve a Long-Lasting and Far-Reaching Problem. *Simulation & Gaming*, *46*(3–4), 404–429. http://doi.org/10.1177/1046878115580412

Beall, A., & Ford, A. (2011). Learning from System Dynamics Simulations.

Black, L. (2013). When visuals are boundary objects in system dynamics work. *System Dynamics Review*, *29*(2), 70–86. http://doi.org/10.1002/sdr.1496

Black, L., & Andersen, D. (2012). Using Visual Representations as Boundary Objects to Resolve Conflict in Collaborative Model-Building Approaches. *Systems Research and Behavioral Science*, *29*(2), 194–208. http://doi.org/10.1002/sres.2106

Bodily, S. E. (1985). *Modern decision making: a guide to modelling with decision support systems*. McGraw-Hill.

Climate Change Act, § c. 27 (2008).

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. http://doi.org/10.1177/001316446002000104

Davidsen, P. I., & Spector, J. M. (2015). Critical Reflections on System Dynamics and Simulation/Gaming. *Simulation & Gaming*, *46*(3–4), 430–444. http://doi.org/10.1177/1046878115596526

Davies, M., & Oreszczyn, T. (2012). The unintended consequences of decarbonising the built environment: A UK case study. *Energy and Buildings*, *46*, 80–85. http://doi.org/10.1016/j.enbuild.2011.10.043

De Wit, F. R. C., Greer, L. L., & Jehn, K. A. (2012). The paradox of intragroup conflict: a meta-analysis. *The Journal of Applied Psychology*, *97*(2), 360–390. http://doi.org/10.1037/a0024844

Department of Energy and Climate Change. (2012). *The Energy Efficiency Strategy: The Energy Efficiency Opportunity in the UK*. London, U.K.

Eker, S., & Zimmermann, N. (2016). Understanding the Mechanisms behind Fragmentation in the Housing Construction and Retrofit. In *Proceedings of the 34th International Conference of the System Dynamics Society*.

Eker, S., Zimmermann, N., & Carnohan, S. A. (in preparation). Integrated Decision-making in Housing, Energy and Wellbeing: A Participatory System Dynamics Approach.

Eskinasi, M., & Rouwette, E. (2004). Simulating the urban transformation process in the Haaglanden region, the Netherlands. Presented at the 22nd International System Dynamics Conference.

Eskinasi, M., Rouwette, E., & Vennix, J. A. M. (2009). Simulating urban transformation in Haaglanden, the Netherlands. *System Dynamics Review*, *25*(3), 182–206. http://doi.org/10.1002/sdr.423

EU referendum - GOV.UK. (2016, June). [Government]. Retrieved July 23, 2016, from https://www.gov.uk/government/topical-events/eu-referendum

Folger, J. P., Hewes, D., & Poole, M. S. (1984). Progress in the communication sciences (Vol. 4). Ablex, Norwood, NJ.

Ford, A. (1996). Testing the snake river explorer. *System Dynamics Review*, *12*(4), 305–329. http://doi.org/10.1002/(SICI)1099-1727(199624)12:4<305::AID-SDR110>3.0.CO;2-4
Ford, A. (2010). *Modeling the environment*. Washington, DC: Island Press.
Forrester, J. W. (1961). *Industrial dynamics*. Wiley.
Forrester, J. W. (1969). URBAN DYNAMICS. Retrieved from http://trid.trb.org/view.aspx?id=131369
Forrester, J. W. (1971a). Counterintuitive behavior of social systems. *Theory and Decision*, *2*(2), 109–140. http://doi.org/10.1007/BF00148991
Forrester, J. W. (1971b). *World dynamics*. Wright-Allen Press.
Francis, J., Giles-Corti, B., Wood, L., & Knuiman, M. (2012). Creating sense of community: The role of public space. *Journal of Environmental Psychology*, *32*(4), 401–409. http://doi.org/10.1016/j.jenvp.2012.07.002
Franco, L. A., & Montibeller, G. (2010). Facilitated modelling in operational research. *European Journal of Operational Research*, *205*(3), 489–500. http://doi.org/10.1016/j.ejor.2009.09.030
Franco, L. A., & Rouwette, E. (2011). Decision development in facilitated modelling workshops. *European Journal of Operational Research*, *212*(1), 164–178. http://doi.org/10.1016/j.ejor.2011.01.039
Frumkin, H. (2003). Healthy Places: Exploring the Evidence. *American Journal of Public Health*, *93*(9), 1451–1456.
Ghaffarzadegan, N., Lyneis, J., & Richardson, G. P. (2010). How small system dynamics models can help the public policy process. *System Dynamics Review*, n/a-n/a. http://doi.org/10.1002/sdr.442
Goodwin, P., & Wright, G. (2014). *Decision Analysis for Management Judgment*. John Wiley & Sons.
Größler, A. (2007). System dynamics projects that failed to make an impact. *System Dynamics Review*, *23*(4), 437–452. http://doi.org/10.1002/sdr.365
Guite, H. F., Clark, C., & Ackrill, G. (2006). The impact of the physical and urban environment on mental well-being. *Public Health*, *120*(12), 1117–1126. http://doi.org/10.1016/j.puhe.2006.10.005
HM Government. (2011). *The Carbon Plan: Delivering our low carbon future*. Department of Energy and Climate Change.
Homer, J. B. (1985). Worker burnout: A dynamic model with implications for prevention and control. *System Dynamics Review*, *1*(1), 42–62. http://doi.org/10.1002/sdr.4260010105
Hovmand, P. S. (2014). *Community based system dynamics*. New York: Springer.
Kopainsky, B., Alessi, S. M., Pedercini, M., & Davidsen, P. I. (2015). Effect of Prior Exploration as an Instructional Strategy for System Dynamics. *Simulation & Gaming*, *46*(3–4), 293–321. http://doi.org/10.1177/1046878113517536
Lane, D. C. (1995). On a Resurgence of Management Simulations and Games. *The Journal of the Operational Research Society*, *46*(5), 604–625. http://doi.org/10.2307/2584534
Lock, M. M., & Nguyen, V.-K. (2010). *An anthropology of biomedicine*. Chichester, West Sussex ; Malden, MA: Wiley-Blackwell.
Luna-Reyes, L. F., Martinez-Moyano, I. J., Pardo, T. A., Cresswell, A. M., Andersen, D., & Richardson, G. P. (2006). Anatomy of a group model-building intervention: building dynamic theory from case study research. *System Dynamics Review*, *22*(4), 291–320. http://doi.org/10.1002/sdr.349
Macmillan, A., Davies, M., Shrubsole, C., Luxford, N., May, N., Chiu, L. F., … Chalabi, Z. (2016). Integrated decision-making about housing, energy and wellbeing: a qualitative system dynamics model. *Environmental Health*.
Maier, F. H., & Größler, A. (2000). What are we talking about?—A taxonomy of computer simulations to support learning. *System Dynamics Review*, *16*(2), 135–148. http://doi.org/10.1002/1099-1727(200022)16:2<135::AID-SDR193>3.0.CO;2-P
Martin, L. J., Hathaway, G., Isbester, K., Mirali, S., Acland, E. L., Niederstrasser, N., … Mogil, J. S. (2015). Reducing Social Stress Elicits Emotional Contagion of Pain in Mouse and Human Strangers. *Current Biology*, *25*(3), 326–332. http://doi.org/10.1016/j.cub.2014.11.028

Meadows, D. (1989). Gaming to Implement System Dynamics Models. In P. D. P. M. Milling & P. D. E. O. K. Zahn (Eds.), *Computer-Based Management of Complex Systems* (pp. 635–640). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-74946-9_69

Morecroft, J. D. W. (1988). System dynamics and microworlds for policymakers. *European Journal of Operational Research*, *35*(3), 301–320. http://doi.org/10.1016/0377-2217(88)90221-4

Objectives | European Master in System Dynamics. (2014). Retrieved from http://europeansystemdynamics.eu/sample-page/objectives/

Pavlov, O. V., Saeed, K., & Robinson, L. W. (2015). Improving Instructional Simulation With Structural Debriefing. *Simulation & Gaming*, 1046878114567980. http://doi.org/10.1177/1046878114567980

Prochorskaite, A., Couch, C., Malys, N., & Maliene, V. (2016). Housing Stakeholder Preferences for the "Soft" Features of Sustainable and Healthy Housing Design in the UK. *International Journal of Environmental Research and Public Health*, *13*(1), 111. http://doi.org/10.3390/ijerph13010111

Richardson, G. P. (2013). Concept models in group model building. *System Dynamics Review*, *29*(1), 42–55. http://doi.org/10.1002/sdr.1487

Rouwette, E., Korzilius, H., Vennix, J. A. M., & Jacobs, E. (2011). Modeling as persuasion: the impact of group model building on attitudes and behavior. *System Dynamics Review*, *27*(1), 1–21. http://doi.org/10.1002/sdr.441

Rouwette, E., Vennix, J. A. M., & Mullekom, T. van. (2002). Group model building effectiveness: a review of assessment studies. *System Dynamics Review*, *18*(1), 5–45. http://doi.org/10.1002/sdr.229

Ruth, M. (2015). *Handbook of Research methods and Applications in Environmental Studies*. Edward Elgar Publishing.

Ruud, M., & Baakken, B. (2003). Development of multiplayer games through group modelling. In *Proceedings of the 21st International Conference of the System Dynamics Society*. New York City, USA.

Shrubsole, C., Macmillan, A., Davies, M., & May, N. (2014). 100 Unintended consequences of policies to improve the energy efficiency of the UK housing stock. *Indoor and Built Environment*, *23*(3), 340–352. http://doi.org/10.1177/1420326X14524586

Snabe, B., & Größler, A. (2006). System dynamics modelling for strategy implementation—case study and issues. *Systems Research and Behavioral Science*, *23*(4), 467–481. http://doi.org/10.1002/sres.773

South, J. (2015). *A guide to community-centered approaches for health and wellbeing*. Public Health England.

Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387–420. http://doi.org/10.1177/030631289019003001

Stave, K. (2010). Participatory System Dynamics Modeling for Sustainable Environmental Management: Observations from Four Cases. *Sustainability*, *2*(9), 2762–2784. http://doi.org/10.3390/su2092762

Stave, K., Beck, A., & Galvan, C. (2014). Improving Learners' Understanding of Environmental Accumulations through Simulation. *Simulation & Gaming*, 1046878114531764. http://doi.org/10.1177/1046878114531764

Sterman, J. (1992). Teaching Takes Off. *OR/MS Today*, *35*(3), 40–44.

Sterman, J. (2000). *Business dynamics : systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill.

Sterman, J., Fiddaman, T., Franck, T., Jones, A., McCauley, S., Rice, P., … Siegel, L. (2012). Climate interactive: the C-ROADS climate policy model: Climate interactive. *System Dynamics Review*, *28*(3), 295–305. http://doi.org/10.1002/sdr.1474

Sterman, J., Fiddaman, T., Franck, T., Jones, A., McCauley, S., Rice, P., … Siegel, L. (2013). Management flight simulators to support climate negotiations. *Environmental Modelling & Software*, *44*, 122–135. http://doi.org/10.1016/j.envsoft.2012.06.004

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.)*, *185*(4157), 1124–1131. http://doi.org/10.1126/science.185.4157.1124

UCL Innovation and Enterprise. (2016). Retrieved August 18, 2016, from https://www.ucl.ac.uk/enterprise

Van Daalen, C., Schaffernicht, M., & Mayer, I. (2014). System Dynamics and Serious Games. Presented at the 32nd International Conference of the System Dynamics Society, Delft, The Netherlands, 20-24 July 2014; Authors version, System Dynamics Society.

Vennix, J. A. M. (1996). *Group Model Building: Facilitating Team Learning Using System Dynamics* (1 edition). Chichester ; New York: Wiley.

Vennix, J. A. M. (1999). Group model-building: tackling messy problems. *System Dynamics Review*, *15*(4), 379–401. http://doi.org/10.1002/(SICI)1099-1727(199924)15:4<379::AID-SDR179>3.0.CO;2-E

Vennix, J. A. M., & Rouwette, E. (2000). Group model building: what does the client think about it now? In *Proceedings of the 18th International Conference of the System Dynamics Society* (pp. 208–209). Bergen, Norway.

Vennix, J. A. M., Scheper, W., & Willems, R. (1993). Group Model-Budding: What Does the Client Think of It? In *Proceedings of the 1993 International System Dynamics Conference*.

Villanueva, K., Badland, H., Hooper, P., Koohsari, M. J., Mavoa, S., Davern, M., … Giles-Corti, B. (2015). Developing indicators of public open space to promote health and wellbeing in communities. *Applied Geography*, *57*, 112–119. http://doi.org/10.1016/j.apgeog.2014.12.003

Wijnen, F. M., Mulder, Y. G., Alessi, S. M., & Bollen, L. (2016). The potential of learning from erroneous models: comparing three types of model instruction. *System Dynamics Review*, n/a-n/a. http://doi.org/10.1002/sdr.1546

Wikibooks contributors. (2016, June). Scriptapedia - Wikibooks, open books for an open world. Retrieved August 11, 2016, from https://en.wikibooks.org/wiki/Scriptapedia

World Health Organization (Ed.). (2012). *Addressing the social determinants of health: the urban dimension and the role of local government*. Copenhagen, Denmark: World Health Organization Regional Office for Europe.

Zimmermann, N., Black, L., Shrubsole, C., & Davies, M. (2015). Meaning-making in the process of participatory system dynamics research. In *33rd Meeting of the International System Dynamics Society*.

Appendix A: Observations of Small Workshops

Workshop 1: Industry

Overview of Process and Outcomes

The first workshop was held with two academic members of the UCL IEDE, new to the process, and one business/industry representative. Both of the UCL- affiliated participants were invited the morning of the workshop due to last-minute cancellations by previously confirmed attendees. However, they had experience relative to the industry topic area as each had several years of experience in industry previously. Participants were welcomed to the session, and given an overview of the previous work in the HEW project in order to situate need for the day's workshop relative to previous efforts.

After this brief introduction, the participants' focus was directed towards a wall in the room where a grouping of variables had been placed. These variables were chosen based on a coding analysis of previous industry-affiliated stakeholder interviews. The participants were asked to consider the variables and to suggest new variables they think are important, but missing. The facilitator reflected back variable names asking for clarification when necessary. As one UCL stakeholder was late in arriving, only two took part in the generation of additional variables. In total, seven new variables were added, shown in the table below.

Table 14. Variable elicitation from industry workshop (new variables in bold).

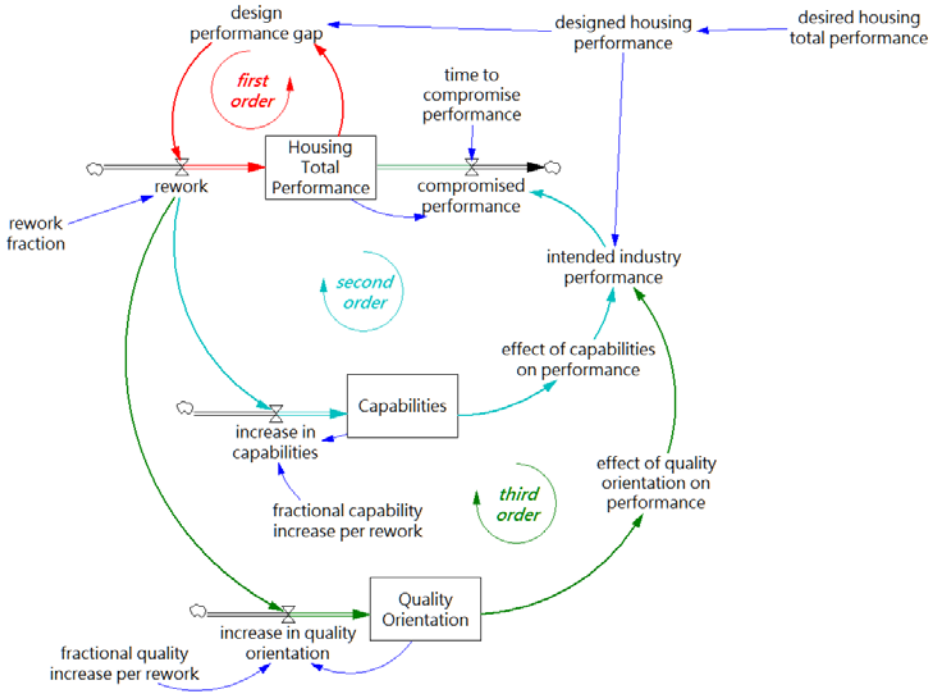
Variable	*Votes*
Building regulations	8
Monitoring by developers	7
Occupants' desired performance	6
Mental and emotional wellbeing	3
Communication between designer and constructor	3
Experience of workforce	3
Training the workforce	3
Liability	3
Integration of scheduling	3
Occupants' valuing of performance	3
Type of finance	3
Rework	2
Community connection	1
Carbon emissions from the housing stock	1
Energy efficiency	1
Consistency in the housing polices	1
Total housing performance	1
Physical wellbeing	1
Constraints by component manufacturers	1
Innovation	1

Capabilities	1
Quality orientation	1
Availability of finance	1
Line of communication between finance and constructing	1
Performance as designed	0
Third space performance	0
Land allocation	0
Developers' share in the housing market	0
Diversity of workforce	0
Training managers and board members	0

New variables were usually accompanied by a brief description of the participants reasoning. In most cases, these descriptions related to other variables that had been suggested by the interview results. Once the participants had no new variables to add, the dots script was carried out (still with two participants), with each given 20 votes to assign priority to the variables. The facilitator briefly summarized the results back to the participants and highlighted the connections that had been made to variables which would appear in the concept model in the following exercise.

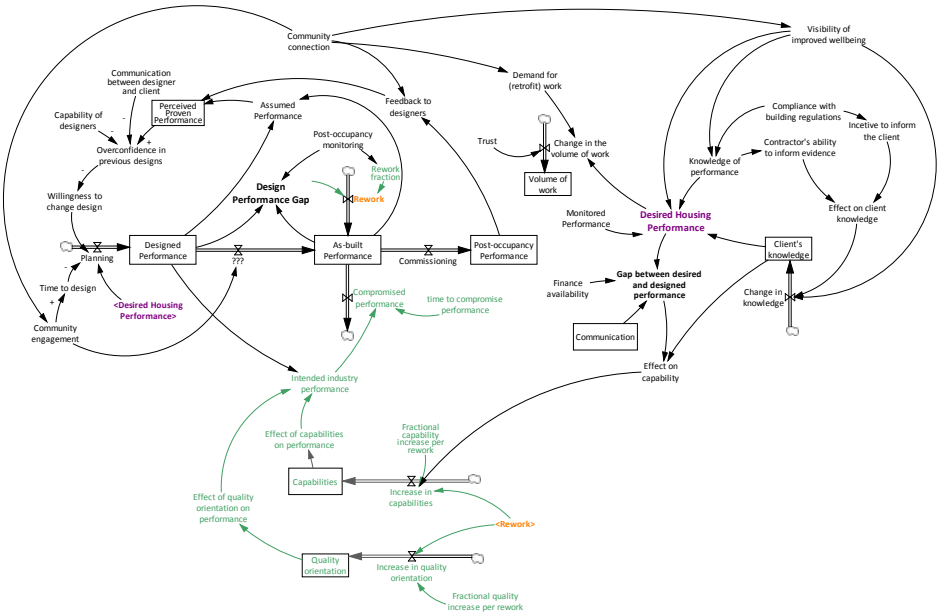
Next, the model was unfolded gradually and the dynamic hypothesis explained. Participants were encouraged to point out errors and the facilitator indicated specifically that the model was not intended to be comprehensive, but rather serve as a representation of the most important problems being faced in the industry sector as determined from previous interview data. Participants agreed with the dynamic hypothesis and were engaged in understanding the structure-behavior relationship as the unfolding gave way to simulation and sensitivity analysis. They focused on the "total housing performance" portion of the structure, and sought further explanation of the flows influencing this stock. Participants were asked if they needed further clarification before being directed on to a glass wall in the room where the model structure had been drawn (and covered) before the workshop start.

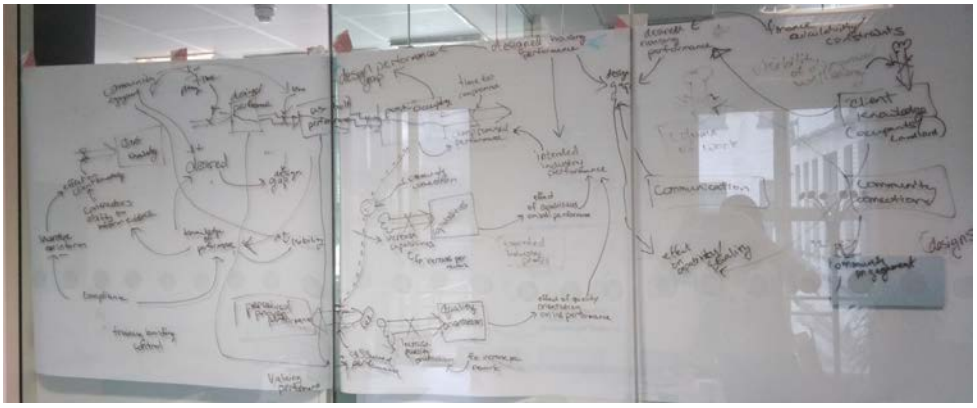
Structure elicitation then began in earnest, and the concept model proved effective at stimulating discussion and raising the level of cognitive conflict. Approximately 30 minutes into the structure elicitation, the second UCL-affiliated stakeholder arrived, and was given a brief summary of the state of the process by the facilitator. There was a good level of participation among the group, and new structure readily emerged around the basis of the concept model. The mix of experience among the group lead to a diversity of ideas and connections. Examples from personal experience was often used as a medium for communication, requiring the facilitator to relate the discussion back to model structure. There was occasional disagreement about the formulation of some of the causal links, but this appeared as task conflict and related to the model. Relationship conflict did not appear. Process conflict was also absent, and participants did not ask questions about the need for the tasks at hand.



As model structure was drawn on the wall by one facilitator, another followed along with Vensim® software copying the developments. Both the modeling and process facilitator also interjected and asked clarifying questions or made suggestions. As the structure elicitation segment came to a close, the group was asked how they could relate the structure to the other topic areas that would be the focus of the workshops to follow. Community, community connection and community engagement were highlighted as having many connections. Time did not allow for the full exploration of policy. However, throughout the structure elicitation, policy suggestions were made by the participants in relation to the emerging S&F diagram. The disconfirmatory interview strategy was applied and seen to be effective in the group setting, perhaps due to the small size.

The result of the structure elicitation is seen in the figures below.





A brief summary was given to the participants regarding the days' progress and the next steps of the HEW project and questionnaires were distributed. However, the final discussion and next steps were extremely brief because two stakeholders had to leave early. At this time the late-arriving stakeholder was asked to include his own dots to add to the votes of the others.

This was followed by a debrief session among the facilitator team. A general analysis of the workshop took place and three specific areas were addressed, those were: what went well, areas of concern and areas for improvement for the next workshop.

Benefits

- Nici found substitute participants quickly
- We could elicit new structure we did not consider/know before
- Shane paid attention to have a balanced participation and explicitly asked for each participant's opinion
- Cooperation among facilitators
- Output
- No stakeholder dominated
- Participation, engagement, engagement increased partially
- Participants understood concept model
- Small group size
- The substitute participants turned out to be useful

Concerns

- The structure was only in terms of relations. We could get more, at least about the directions of links etc., and we could 'converge' more
- We only had one 'real' participant (→ What impression does this make on the participant?)
- No sheets of paper for the variable elicitation session
- Nici could often not see the board and thus not make comments to what was going on
- End of meeting deflated
- Russell got quieter in the end

POST-WORKSHOP:

We aimed to send a cleaned version of the model developed in the session to the participants in a few days following the session. However, we decided to send a 'running' model with preliminary results, and this took longer time due to the Easter holiday. Eventually, the model was shared with participants in a visual form (PowerPoint file). The file contains brief explanations about our follow-up assumptions, and preliminary simulation results. The participants were asked to comment on this model if they have any concerns. No response has been received so far.

Workshop 2: Community

Overview of Process

Technical difficulties lead to some deviations from the designed schedule. In particular, the lack of a working projector made the presentation of previous HEW work difficult. After a brief round of introductions, the five participants were asked to gather around a laptop and a shortened version of the presentation was made. Next, a transition to the dots exercise took place. Participants were eager to share stories and added many variables while discussing the meaning of the variables already suggested by the modelling team. They recognized and mentioned the links between the variables already in this stage. Some individuals in this group tended to dominate discussions, speaking for extended periods which limited the contributions of other participants. This combined with the start-up delay from technical problems and led to this exercise extending well beyond the scheduled time. The facilitator recognized this and moved the pace forward by asking for any final important variables before passing out the sticky-dots. The divergent and convergent stages of this script were readily apparent, with active discussion, storytelling followed by voting. Results can be seen in Table 1. Before transitioning to the concept model the facilitator summarized the themes that emerged.

Table 15. Variable elicitation from community workshop (new variables in bold)

Variable	*Number of Votes*
Physical Mental and Emotional Wellbeing	9
Community Connection	7
Residents' Desired Performance	6
Residents' agency in planning process	6
Housing Affordability	6
Residents' Knowledge of Available Technologies	6
Residents' incentive to get involved	5
Cohesion	5
Land allocated to third spaces	5
Post occupancy evaluation	5
Third spaces	4
Residents' ability to understand planning policy	3
Residents' knowledge of plans	3
Family Connection	3
Length of tenure	3
Building regulations	3
Household Crowding	3
Residents' feedback to policy	2
Ease of parking	2
Land Allocated to Housing	1
Residents' Perceived Performance	1
Feedback to Building Designers	1
Service charges	1
Housing density	1

Ease of access	1
Extent of tenant mix	1
Noise	1
Safety	0

The concept model, seen in Figure 10, was next unfolded, and each of the loops described step by step. There was agreement from the stakeholders regarding the overall scope and content, but the concept model served its purpose of "jump-starting" the discussion. Facilitation of model structure focused a great deal on community connection and third spaces as well as demographic changes and gentrification issues. Participants shared rich stories, and a central variable "use of third spaces" emerged very quickly to form new feedback loops.

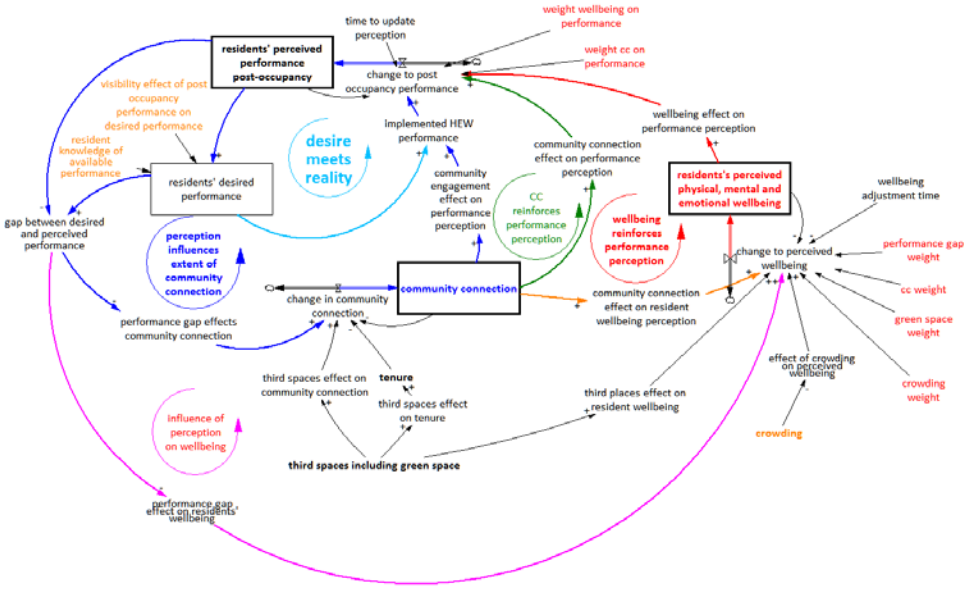


Figure 10. The concept model for the community workshop.

After approximately one hour of this structure elicitation, a break was taken and the modelling team drew several lookup effect axes on the board in order to elicit specific effects from the expert stakeholder group. This included a newly added concept of effect of community connection on loneliness. This was intended as a convergent exercise to gain consensus and to discuss validity of portions of the model. To a degree, this was accomplished, especially for the effect of loneliness on wellbeing for which there was wide agreement. Some participants confirmed links via storytelling. However, this did not occur for the effect of wellbeing on perceived performance, which showed that participants saw much stronger connections than this one. The facilitation required frequent referral to

the model and graphs in order to move towards convergence.

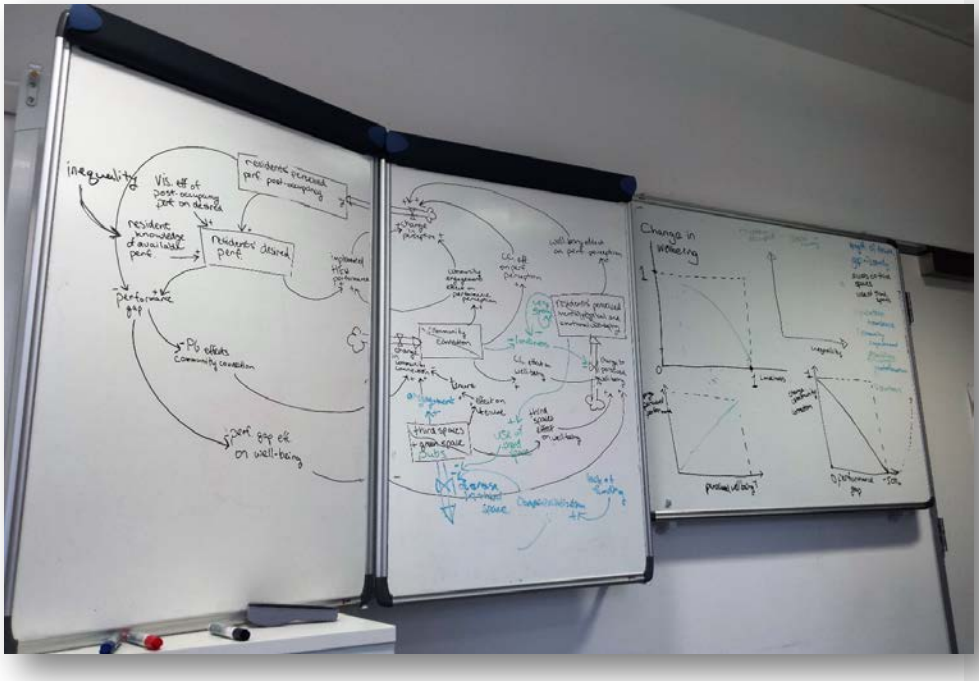


Figure 11. Structure elicitation and lookup effect axes developed in the community workshop.

Finally, during the distribution of questionnaires, the swing weighting technique was used within the context of the factors contributing to a weighted additive function in order to determine the rate of change to perceived wellbeing, a soft variable that is particularly difficult to assess in the context of an SD simulation model. (For more on swing weighting see: (Goodwin & Wright, 2014)).

Some participants clearly contributed more than others, but all stakeholders did contribute. There were no open conflicts between them, and we observed that participants added variations to each other's explanations to more fully capture the situation from their point of view.

Content that was not directly captured in the model structure:

- Third spaces and their use reinforce each other (capacity and capacity utilization)
- Third spaces are not used by the people with income who investors desire as users
- Third spaces are transformed into commercial space
- New and short-term residents do not use third spaces
- Incentives to participate → engaging in community

- Sharing something in common with those around you is an important factor to participate
- Reciprocity, having a stake in the community
- Demographics → tenure → perceived benefit from engaging in the community ← visibility (e.g. of joggers in the park etc.)
- Affordability, health, safety → use of third spaces
- Feeling positive → feeling more able to do something about the environment, agency
- Betting shops (used by Afro-Caribbean and West-Indian men whose other community places do no longer exist)
- Equality
- Diversity → knowledge
- Diversity → - shared experience with the other person → - cohesion
- People rather leave than do something about the community (if they have the option)

What went well?

- Introduced cognitive conflict (divergent) and consensus building (convergent) aspects of GMB successfully.
- New feedback loops were made and participants confirmed others via storytelling, which was related to the model by the facilitator.
- Successful elicitation of weighted additive function via the swing weighting technique
- Concept model jump-started conversation and participants shared rich stories
- Facilitator focused discussion on the model
- Disconfirmatory approach revealed some links to be weak or non-existent in the concept model (i.e. wellbeing effect on residents' perceived housing performance), others were confirmed.
- How we handled two not working projectors
- Collaboration between the facilitators

What went poorly?

- Some people dominated the discussion, though all participants did engage.
- Dots exercise took longer than expected, was not linked well to the structure elicitation phase
- Technical difficulties ate into short workshop time.
- Some participants seemed to struggle with putting stories into structure.
- Lack of connection to other aspects of the broader modelling goals structure, participants were anchored to the model and there was not enough time to spread towards industry and policy concepts.
- A participant who was new to the project did not participate actively.
- We had not finished all preparations in sufficient time before the workshop started

What can we do to improve?

- Come earlier to the location to test all equipment, don't underestimate the time needed to arrange the physical space.
    - 15 minutes before the session: Projector running, concept model on board (in appropriate font size and readability), pens, evaluation sheets, reimbursement forms and participation agreement forms ready
    - 18 hours before the session: concept model finalised
    - 48 hours before the session: good concept model draft

Workshop 3: Policy

Overview of Process:

The session began with a short presentation to remind participants of the work completed in the HEW project thus far. One participant was approximately 20 minutes late, and so missed the introduction to the agenda and aims of the day.

Facilitating the policy group was a more difficult task for every stage of the process. During the introductory dots session participants were directed towards the variable list, and definitions of the variables were given by the facilitator. This kicked off a good group discussion though it seemed to take a while before new variables began to appear along with the discussion. At least half of the group tended to dominate discussions using extended story-telling as a primary method of discussion. The facilitator, when able, would relate the discussion back to the variable list but this was not always successful in identifying new variables that would be important to model. Stakeholders were very engaged, however their tendency to control the floor caused difficulty in pacing the task. From the point the facilitator began to move towards to voting (convergent) point of the task, it took another 15 minutes before a natural break in conversation allowed for the dots to be distributed. One participant suggested that the variables be condensed in order to simplify the process, and this idea was entertained by the facilitator. However, in attempting to facilitate such a 'grouping' more general discussion began to emerge. In response, the facilitator told participants that the voting would serve the same purpose and asserted that it was time to move on. This was followed by a final question to the group regarding any final variables which must be added to the list. Participants did not respond immediately and this silence was accepted as their answer so the voting could begin. Again, during the voting, participants were very engaged in the task, having discussions about the definitions, connections and groupings of the variables as they went. One participant in particular was eager to begin drawing connections between variables, describing a link between learning of policy designers/analysts and competence of policy designers/analysts.

Table 16. Variable elicitation from policy workshop (new variables in bold).

Variable	*Number of Votes*
Silo-ing of expertise of all the building specialists	11
Demand for energy efficiency projects	10
Balance between central and local government	10
Means of delivery	9
Use of output measures	7
Community interest	7
Accountability	7
Resources allocated for HE(W) policies	6
Residents' desired performance	6
Volume of the energy efficiency projects	6
Industry interest	6

Feedback on actions	6
HEW performance in building regulations (standards)	5
Competence of policy designers/analysts	5
Integration of complexity	4
Scope of the responsibility boundary	4
Learning of policy designers/analysts	3
Local authorities' influence on implemented performance	3
Community connection	3
Commitment to a policy direction	2
Local authorities' influence on design performance	2
Breadth/scope of government departments	2
High-performance housing stock	1
Implemented HEW performance	0
Monitoring the policy outcomes	0
Time to get agreement on a policy	0

The group was next taken through a demonstration of the concept model shown below. The facilitator welcomed comments and questions and emphasized that the model is a simple representation, encouraging discussion about what part of the model is wrong. The concept model served its purpose of jump-starting the conversation, with one participant already offering changes towards improving the model before demonstration had finished. Once the loops had been described the participants were directed to the whiteboard and disconfirmatory questions posed regarding two key loops: policy scrapping and learning. Participants confirmed the existence of these loops and their importance for policy making. Later in the session participants would help develop table functions to describe the action of these loops, to lend validity to future modeling efforts.

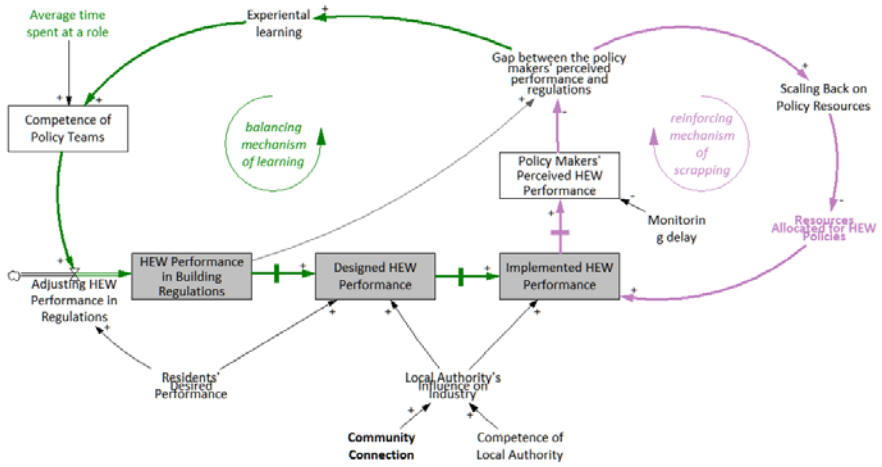


Figure 12. Concept model presented in the policy workshop.

The model served well as a boundary object but creation of new structure in a coherent manner was more difficult than in previous sessions and engaged participants manifested in different ways. One participant (was engaged with trying to close loops, and even offered to bring the discussion back to the model structure on his own accord. At several points in the workshop he jumped out of his seat and came up to the board to aid in his description of loops he saw as critical. One participant had a background in SD model building for housing policy. He sometimes helped direct the group back toward the method offering suggestions in much the same way a facilitator might. This added a unique element to the workshop that was absent in previous workshops.

This was an interesting outcome given that most participants had familiarity with the CLD method from previous workshops or a background in SD modeling. The facilitator directed participants back to the model structure, but not all were willing to use the model as a means to expand their ideas. Story telling was a more comfortable method for expression, either using hypotheticals or specific examples from past policy successes/failures. The size of the group had a noticeable impact. At eight participants, it was the largest of the three small workshops. This added to the difficulty in keeping the discussion focused on the model, but also aided in bringing beneficial convergent thinking to the fore. A general observation was that they were able to capture loops, if the number of associated variables and connections was less extensive. For example a reinforcing loop that was easily recognized regarding the tendency of policy to become more centralized thereby reduce the number of stakeholders involved in decision making, increasing centralization. Yet they were not thinking much towards closing the loops that involve a longer chain of variables.

By the end of the structure elicitation process there was certainly a noticeable converging of thoughts among the group. Participants commented on each other's narratives in general agreement, using phrases like "that's exactly the problem", rather than waiting for a pause in another's story so that they could begin their own. When newly suggested links among community connection, local authority's influence on industry and the planning process emerged, a participant representing a local authority of a London borough said to the group, "that is the ideal situation." Further convergence came about in the group elicitation of table functions for the learning and policy scrapping loops, revisited from the original concept model structure. While the policy scrapping function was agreed upon fairly quickly by the group, the learning function was more difficult and participants had different views. Since no consensus was achieved for that table function, participants were asked to draw the axis and plot the graph on their questionnaires.

To close the session, the outcomes of the day were briefly reflected upon and the bulls-eye diagram showing the emphasis areas of the modeling effort thus far was explained. Participants were then guided through a swing-weighting exercise for factors affecting the variable "implemented HEW performance" comparing policy, community, industry and all others influences. There was some confusion among participants about how to carry out the exercise. They showed strong awareness of scenarios and wondered whether they should prioritize based on their ideal future of the influences or the current situation. Even though the session went slightly over time, most participants took time to fill out questionnaires, except for one who had to leave for another appointment.

Appendix B: Coding

The Coding Manual

Table 17. The coding manual developed to analyze the transcribed audio data.

Research Objective	Method and Data Source	Positive Codes	Negative Codes
Learning and insight	CICC and audio data	• Inquisitive statements regarding counterintuitive result • Statements of surprise • Describing feedback • Explicit statements of understanding	• Rejecting counterintuitive results without inquiry • Describing relations as linear • Explicit statements of not understanding.
Boundary object in game	Stage 4 of (Black & Andersen, 2012), audio data, CICC - commitment	• Modification of visual element • Participants make reference to visual element to construct an argument • Statements about tool usefulness, in present and in future applications. • Expressed interest to share results or ideas from workshop with others.	• Participants avoid discussing the visual element • Causal descriptions that don't relate to the visual object • Expressed frustration with visual element (e.g. inability to change as desired) • Expressing doubts about model uses in present or future. • Expressed difficulty using tool

Boundary objects in GMB workshops	Stage 1-3 of (Black & Andersen, 2012), audio data	• Modification of visual element • Participants make reference to visual element to construct an argument • Statements about visual element's usefulness. • Discussion of interrelationships and dependencies • Addition of new variables	• No additional variables proposed • Viewing variables as isolated or not connected • Rejection of terms used by others. • Participants avoid discussing the visual element • Arguments that don't relate to the visual object • Expressed frustration with visual element (e.g. inability to change as desired)
Fragmentation	CICC – consensus & communication	• Converging of opinions, agreement, and accommodation of other points of view. • Common language use • Participants share speaking turns.	• Disagreement, unwillingness to accommodate other points of view • Use of different vocabulary • Participants give multiple lengthy descriptions
Consideration of multiple objectives	No particular method, audio data	• Positive statements pertaining to community, third spaces, mental or physical health or wellbeing. • Agreeing on definitions for soft variables • Expressed desired to have higher performance on multiple indicators and soft variables	• Negative statements pertaining to community, third spaces, mental or physical health or wellbeing. • Disagreement about definitions for soft variables • Concern with one indicator only

The bottom half of every page was coded according to the above table by first looking for negative indications, followed by a search for positive indications to mitigate confirmation bias. For each speaking turn, multiple phrases giving the same indication, in the same direction were included. For example, if in a single speaking turn a participant made a positive statement regarding community such as “we should invest in parks” followed by a positive statements pertaining to wellbeing such as “I think people’s health is important, too” would be coded as two positive indications.

Section of Transcripts used for Training

Game Workshop: Table 2, Trial Driven

First Speaker: F1
Second Speaker: P1
Third Speaker: P2
Fourth Speaker: P3
Fifth Speaker: P4
Sixth speaker: F2

--00:00—
SE: Sorry could you repeat the question?
P1: The energy efficiency performance, is it modeled based on change in climate as well? So thinking about how to build the performance in terms of energy for cooling or is it just energy for keeping homes warm?
F1: Well we can't say all, we don't go into this detail so its not separate like cooling and heating, but it in general refers to energy consumption which can be used either for cooling or heating.
P2: We also don't know what our policy priorities are so if we're looking at where we're starting from, we're at .65 on the energy efficiency for the housing stock, and we're at .58 or something for overall wellbeing performance on houses, buildings. So you could say, are we going to fix the things that are low value to begin with.
F1: I think that is something you can decide on altogether
P2: Well I'm asking the question, not of you, I'm asking the question for our group
P3: I mean that might be a useful thing to do, to sort of set some policy objectives and try and see maybe one maximizing wellbeing and another one maximizing….
P1: So what if you're both maximizing wellbeing and addressing inequality, what if those are your two priorities, which is my understanding is two of the organization priorities.
P2: So this is the built environment budget though?
P1: Quite difficult, I think that comes to the heart of why this is a very useful exercise and useful tool.
P3: Yeah, we don't tend to think of like a holistic view, we don't take a minute to sort of aggregate cost benefit and cost, say of social division, health care and the cost of housing but if we were to put an overall wrapper over that and optimize the cost, so you actually want fewer people seeking medical help, y'know apart from wellbeing, I think that's a given. Um, which is the dominant effect in policy were you able to put a wrapper over so to say, social care and health service?

P1: Yeah

P2: The started in 2000 they started invested health funding into improvement of existing homes where they were very poor quality.
P3: Yes, but they haven't fully, if you look at the impact assessments I dunno what the latest detail conversation shows but they haven't been, they sort of started costing the health care benefits but not fully integrated those into the…
P2: But also we haven't got the health care budget, we've got the long term built environment budget which is different. So what's our priority? Politically what's our priority?

P1: Yeah
P3: Well presumably it's the, the value of the stock as well isn't it. So you're actually appreciating that.
P1: yeah
P2: But we don't have a measure of that, so we can't take account of that. I mean we can say that we think we should, and we can make assumptions about that ourselves and say that they ought to build that into the model, but I think that's a good point.
P3: Yeah, its a missing factor isn't it, it would be in your mind if you were doing it
P1: Well partly, if you were thinking as the mayor and as the GLA you'd be thinking, you might be thinking as the GLA about those stock owners that need to actually be responsible for their stock. You've already got refit programs, perhaps if you do some of the work with the communal spaces, given its such a broad definition, you might be able to increase the commercial success of areas in London and that might be on the mind as well.
P3: So touching on your use of shared space, the common place where money exchange goes on its sort of the marketplace and social as well, so.
P2: I suppose the thing that troubles me slightly is that we are experimenting with an unknown structure here, the structure of the model.
F1: Uh but you have your own mental model I guess

P2: In the real world you don't have access to the model either do you, so it is a sort of thing you can play with
F1: Yes, so you use your own mental model of how these are connected to each other
P2: But we can change the model, we can put the financial question in
F1: Later, not now
P2: Maybe observing those things that are missing might be useful as well
P3: So, we're the built environment people, we're not the health care people and we're not the social care people.
P2: No, but I think it is, in terms of what are the impacts of spending in that area across a whole range of policy areas.

P3: Agreed, so this is devil's advocate. I am very concerned that 50% of Co2 emissions come from the built environment 27% from UK housing, so that is very high priority on my list. I don't want to make community connection worse, and I would like to make it better if I can, but actually the amount of difference I can make is probably fairly small a lot of it is to do with social connections.
P2: Yes, there are other factors, that's the other thing isn't it. So this isn't a highly sensitive driver for those
P3: For that one I think and the actual wellbeing of residents I think exactly the same. You can make the wellbeing very bad with poor quality buildings but then once they get to a certain point its very hard for the buildings themselves to make people have a much higher level of wellbeing its then down to the connections that they've got with groups of people, the support networks they've got the friendships they've got. So I don't think those two are that sensitive to the level of investment you put into the buildings, again devils advocate, so…
P2: Well particularly in energy efficiency, if you were somehow refurbing them so they were nicer places

to be in terms of overall
P3: Yeah changing in the stage. So I think we should focus more on the first two because we're the built environment fund, and that's where our money should go. And I think they are much more sensitive to where we put our money.
P1: Can I tell you what the built environment people would do, is that they would put nearly all of it in the number one, energy efficiency, and ultimately, legally the UK is required to report on the carbon emissions reductions, that's a legal target, irrespective of any changes anew that's an international target we agreed to and we're likely to ratify COP21. So I guess I would be thinking of that, and if I wasn't thinking of what the mayor wanted, because I could probably dress it up either way as health and indoor quality, if I was literally thinking in the mindset of my ex-colleagues, the reality would be I would pump nearly all of it into energy efficiency of the housing stock.
P2: So then I get back to my actual experience, which is, you have to monitor that to make sure you get it, otherwise the quality of the workmanship, you don't actually get an insight into the behaviors of how people use the property um, and you can't feed anything back to secure your investment, in non-financial terms.
P4: But do you need 20% or 10%?
P2: We don't know the sensitivity because we don't know the structure of the model is, so its. My starting position was, if you're viewing it as in terms of its importance rather than the money you're putting in I would say starting from 50 50 point of view and maybe trading 10% on the way so maybe 60 40 and that might be an interesting test.
P3: But again, I come back to my argument, that its not just about that we are only the built environment and there are other people investing in other areas so we need to work out what our priorities are from the built environment side.
P4: I think what we are talking our way towards is another run which is obviously very heavily weighted towards the energy efficiency side I think surely, I'm not quite sure I understood your point just then because surely if you are going to invest very heavily on the energy efficiency side saying yknow that this is the carbon one, where we'll have the most impact. And surely we need to invest in the monitoring to make sure that that has the..
P3: That you need to invest 20% or 10% is what I'm asking, I don't know how sensitive or how much money that is in real terms and if you put in say 100,000,000
P4: Well it think we don't know do we, I mean..
P2: You might spend I don't know sort of, in monetary units, rather than focus units what, 50 units for energy efficiency that's what 20 or 10 for marching cause you get a lot for you money in terms of functionality.

P3: Exactly
P2: But that's not the way it seems to be scaled here
P4: Its difficult to tell though isn't it

P3: it's the exchange rate for effect isn't it
P2: So if I'm thinking of a hundred million going into housing, and 10 million going into monitoring. With 10 million you can do a lot of monitoring and that's only 10% of the cost of the built environment

intervention.
P3: Bu I think what we're saying is we don't know what the assumptions are about how spend turns into effort.
P2: So I'd argue for going up on energy efficiency and going down on monitoring
P4: But down on communal spaces as well?
P2: Possibly, yes for the next run of the model, just because if you think we need to do is use the model to check the sensitivity
P3: Yeah, that would seem the interesting thing to me to do is to cut down the communal spaces given the conversation we had 5 minutes ago
F1: So more in communal spaces?
P3: No, a lot less in communal spaces
F1: How much?
P3: Can we halve that?
P4: Well I disagree in terms of the monitoring vs. the energy efficiency investment in terms of emphasis rather than the cost. I certainly wouldn't expect to pay as much for the monitoring as the energy efficiency but in these terms if we turn it right down I think that we
P3: But this is monetary units
P1: I just wanted to say that I think what we're about to do now is the run that is a bit more likened to business as usual, its not something we as individuals bringing expertise to the table think, but what we're doing is doing a run to see what we perceive a kind of BAU approach right now might be to actually understand the difference. We started off with on that was much more reflective of our thoughts of what should happen so now we're doing a sort of BAU approach to see what will actually show up.
P2: So on that basis, what is the business as usal spend on monitoring in programs.
P1: it's a lot less.
P2: yeah, yah.
P1: There is monitoring, you can have like 1 in 10 and that is certainly a way to get the money into the… 10% is really the housing grant that supposedly get tested. But that wouldn't be 10% of the 10000 it would be the actual money for it would be.
P3: Are we gonna go for 70?
F1: For energy efficiency
P3: Well don't write it down yet, 70, 15, 15 or 70 20 10
P2: Well just try that because if we keep on constant we can learn a bit more than if we change
P3: We've only got three runs though so we can't test…

P2: No, no. But its in the right direction isn't it
P4: So what are we saying?
P2: Has someone written down or have we got a print out or a not of what the results were?
F1: They are recorded here so we will keep seeing it.
P4: I was gonna say my interest would be to cut the communal if we were looking at 70, 20, 10. Would be to cut the communal spaces budget further just to try and get more of a sense of…
F1: You suggest zero or 1500 for communal spaces?

P4: Well say, 15 is half of what we put in last time.
F1: So we go for 70, 15, 15? So, what do you expect now as an outcome of this decision, how these graphs will look like you think?
P2: That's an interesting question, because 90% of the energy efficiency is quite good. I don't think its asymptotic there. So do we expect to see a higher outcome there basically because that's sort of the intervention we're going for?
P1: If you actually think about London's, I don't know how much this has got real stuff in it anyway, but if you think about the housing stock both, you can do the hard to treat property but beyond that you're not actually making..
P3: No, you're in the tail aren't you.
P1: Especially if we aren't considering cooling, which I think we might not be.
P2: Well I think that 90% is a quite a good outcome, if that's, I don't know what the exchange rate is then necessarily if the maximum we can achieve is unity. What kind of asymptotic sort of … 87 or 85

P3: Okay, so we're testing that out by raising it to 70%
P2: My guess is we won't see a tremendous increase in that.
P3: We're testing out if you only invest 15% in communal spaces does that have a negative impact on any of these things which is what we're trying to avoid.
P2: I doubt it somehow
P3: And the monitoring was reduced by a little bit but not by much
P1: I think it should decrease wellbeing of residents but not by much.
F1: Okay so if this is the agreement

All: Yes
--15:00--
F1: So now the green line is the new result, new simulation.
P2: So it makes nearly no difference
P3: So it's an extremely insensitive intervention
P4: I mean the only one that changed with any size was communal spaces.
F1: That actually went down
P1: It was a negative difference, of course.
P2: So you might say that wasn't a good investment in terms of outcome even though our criterion was to improve the building performance overall.
F1: Well housing performance doesn't look so bad but people aren't so happy.
P3: But then the difference is so little in all the cases
P2: Well its diminishing returns isn't it?
P1: Well other than maybe communal spaces
P3: So the question is what's the number to invest in energy efficiency to get you so we know we can get the change.
P2: 50% actually didn't make a, it made a step change but we, so if you can use sort of. I don't know what the correct terms are mathematical. So plus delta, as a result of 2000 units increase we've got a very small plus delta. And we might say as a result of a 2000 unit decrease we'd have a very small minus

Appendix C: Questionnaires and Game Supplements

Questionnaires:

GMB Workshop

The following is an example of the CICC questionnaire used in the process. For each workshop the section pertaining to validity was modified to reflect the different concept models used. An extended questionnaire used for the game workshop follows.

Dear participant

This questionnaire evaluates the use of Group Model Building (GMB) in the HEW project.

We politely invite you to answer these questions as best as you can. The results of this questionnaire will be used to improve both the procedure of Group Model-Building used and the resultant model.

All information will be treated confidentially if you declare so.

Thank you for your co-operation.

Which of the following best fits your role (please circle one):

Policy analyst / strategy developer

Other local or central government employee

Member of a non-government organisation

Public health professional

Business /industy owner /employee

Politician

Academic

Other: .....

Result of today's modelling process

	agree	some-what agree	neither agree or dis-agree	some-what dis-agree	disagree
1. My insight into the problem has increased due to the modelling process.					
2. The modelling process has given me more insight into the cohesion *between the elements* that compose the problem.					
3. The causal diagrams that were developed were the result of the integration of diverse opinions and ideas of the participants.					
4. If I, with some people from my organisation, were to use the same approach in planning and in dealing with problems, all persons would loyally follow this plan to its natural conclusions.					
5. As a result of the modelling process it is still *unclear* to me what the causes of the problem are that play behind the scenes.					
6. The modelling process aided in the understanding of the opinions of the other participants.					
7. We could *not* reach a consensus.					
8. Our opinions are closer due to the modelling process.					
9. I will uphold the conclusions/findings of these meetings in front of other members of my organisation.					

10. The modelling process has given me *little* insight into the opinions and ideas about the problem of other participants.					
11. Some persons dominated the discussions.					
12. The modelling process has *not* given me insight into the possibilities of addressing the modelled problem.					
13. I will try to convince others in my organisation of the importance of these conclusions.					
14. Using modelling in approaching the problem is efficient.					
15. All in all, I think this meeting was successful.					

Validity of the Group Model Building project

The following questions aim at addressing the usefulness of the model for helping to solve problems and improving thinking about the policy criteria.

	agree	some-what agree	neither agree or dis-agree	some-what dis-agree	disagree
1. It was useful to include the concept of community connection in the model boundary.					
2. It was NOT useful to include the concept of implemented HEW performance in the simulation model					
3. It was useful to include the "gap between policy makers' perceived performance and desired performance" concept in the model.					
4. The structure of the model does not represent the problem as I see it					
5. The simulation results do not capture the problem as I see it					
6. I can describe how the simulation result is created by the structure of the model.					
7. All in all, I gained generic understanding in analysing stock and flow models.					
8. The important issues or problem areas that needed attention were investigated					
9. Most of the variables in the model are well defined, and could be understood by others in my field.					
10. It was NOT useful to include local authorities influence.					
11. It was useful to include the concept of residents' desired performance in the model.					

Are there **additional variables or elements not addressed** in the model that should be added to make it useful to address the problem? Please list these and, if possible, describe the cause and effect relationship given the model structure resulting from today's workshop.

If you compare these meetings, using causal diagrams, with *normal meetings or conferences* in which you discuss *similar problems*, would you say these meetings:

	agree	some-what agree	neither agree or dis-agree	some-what dis-agree	disagree
1. give *more* insight compared with normal meetings?					
2. give insight *more quickly* compared with normal meetings?					
3. result in a *better* communication between participants?					
4. give rise *more quickly* to a shared vision between participants?					
5. give rise to a *better* shared vision between participants?					
6. give rise *more quickly* to commitment of participants?					
7. give rise to *more* commitment of participants?					

Effects of different elements of Group Model Building

The meetings consisted of several aspects which may have contributed in different ways to the overall effect of the meetings. In the following questions you are asked to specify how much an aspect contributed to the overall effect. You can do this by scoring each element on a scale of -5 to +5, in which:

-5 = was of no use whatsoever, obstructed the sessions;

0 = did not obstruct, but was of no use either;

+5 = contributed very much.

	score -5 to +5
The fact that the diagrams were projected/recorded in a way that was visible to everybody.	
The fact that an outsider was accompanying as a 'group facilitator'.	
The opportunity for open and extensive discussion.	
The use of causal diagrams.	
Simulation, using the quantitative model.	
Others,.....	

Please provide your name or any identifier that you will remember in the next workshops. It will help us track the development across workshops. If you want your replies to remain anonymous, please leave this blank.

Identifier:_____

Thank you again for your co-operation.

Game Workshop

Dear participant

We politely invite you to answer these questions as best as you can. The results of this questionnaire will be used to evaluate and improve our work.

All information will be treated confidentially if you declare so.

Thank you for your co-operation.

1. Which of the following best fits your role (please circle one):

Policy analyst / strategy developer

Other local or central government employee

Member of a non-government organisation

Public health professional

Business /industy owner /employee

Politician

Academic

Other: .....

2. In the project we have linked the complex causal loop diagrams, interdisciplinary policy assessment criteria, participatory modelling and an interactive simulation environment. With all that in mind, to **what three specific priority policies** would you recommend to national government to address shared objectives about housing?

1.

2.

3.

3. To what extent has being involved so far in the HEW project changed your thinking about priority policies about housing?

Comments:

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

4. How would you rate the **USEFULNESS** of this kind of modelling approach for supporting more integrated decision-making about housing? Please circle one option:

1	2	3	4
not useful at all			extremely useful

Comments:

5. Thinking about your own role, have you used any of the work we've done so far in the HEW project (circle one)?

- ☐ No, I haven't found the work useful yet, so haven't used it
- ☐ I've found the work useful but haven't yet had an opportunity to apply it
- ☐

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Yes, I've been using the work in my role (please briefly describe how in the box below)

Formatted: English (United States)

How I've used the work so far (e.g. talking about the work with others, using it to assist thinking or debate):

Formatted: English (United States)

6. Result of today's process

	agree	some-what agree	neither agree or dis-agree	some-what dis-agree	disagree
1. My insight into the problem has increased due to the interactive simulation process.					
2. The model and simulation process have given me more insight into the cohesion *between the elements* that compose the problem.					
3. The causal diagrams and model were the result of the integration of diverse opinions and ideas of the participants.					
4. If I, with some people from my organisation, were to use the same approach in planning and in dealing with problems, all persons would loyally follow this plan to its natural conclusions.					
5. As a result of the simulation and modelling process it is still *unclear* to me what the causes of the problem are that play behind the scenes.					
6. The modelling and simulation process aided in the understanding of the opinions of the other participants.					

7. We could ***not*** reach a consensus.					
8. Our opinions are closer due to the modelling and simulation process.					
9. I will uphold the conclusions/findings of these meetings in front of other members of my organisation.					
10. The modelling and simulation process has given me ***little*** insight into the opinions and ideas about the problem of other participants.					
11. Some persons dominated the discussions.					
12. The modelling and simulation process has ***not*** given me insight into the possibilities of addressing the modelled problem.					
13. I will try to convince others in my organisation of the importance of these conclusions.					
14. Using modelling in approaching the problem is efficient.					
15. All in all, I think this meeting was successful.					

7. Validity of the modelling

The following questions aim at addressing the usefulness of the model for helping to solve problems and improving thinking about the policy criteria.

	agree	some-what agree	neither agree or dis-agree	some-what dis-agree	disagree
1. It was useful to include the concept of 'community connection' in the model boundary.					
2. It was NOT useful to include the concept of 'HEW performance of buildings' in the simulation model.					
3. It was useful to include the concept of 'communal spaces' in the model.					
4. Most of the variables in the model are well defined, and could be understood by others in my field.					
5. It was useful to include the concept of 'monitoring' in the model.					
6. The structure of the model does not represent the problem as I see it.					
7. The simulation results do not capture the problem as I see it.					
8. I can describe how the simulation result is created by the structure of the model.					
9. All in all, I gained generic understanding in analysing stock and flow models.					
10. The important issues or problem areas that needed attention were investigated.					
11. It was useful to include the concept of 'energy efficiency' in the model					
13. It was useful to include the concept of 'rework' in the model.					
14. It was useful to include the concept of 'residents' desired HEW performance' in the model.					

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: Font: Not Italic, Font color: Auto, English (United States)

8. Are there **additional variables or elements not addressed** in the model that should be added to make it useful to address the problem? Please list these and, if possible, describe the cause and effect relationship given the model structure resulting from today's workshop.

Comments:

9. If you compare these meetings, using causal diagrams and simulation, with *normal meetings or conferences* in which you discuss *similar problems*, would you say these meetings:

	agree	some-what agree	neither agree or dis-agree	some-what dis-agree	disagre e
1. give *more* insight compared with normal meetings?					
2. give insight *more quickly* compared with normal meetings?					
3. result in a *better* communication between participants?					
4. give rise *more quickly* to a shared vision between participants?					
5. give rise to a *better* shared vision between participants?					
6. give rise *more quickly* to commitment of participants?					
7. give rise to *more* commitment of participants?					

Formatted: English (United States)

10. Effects of different elements of Group Model Building

The meetings consisted of several aspects which may have contributed in different ways to the overall effect of the meetings. In the following questions you are asked to specify how much an aspect contributed to the overall effect. You can do this by scoring each element on a scale of -5 to +5, in which:

-5 = was of no use whatsoever, obstructed the sessions;
0 = did not obstruct, but was of no use either;
+5 = contributed very much.

	score -5 to +5
The fact that the diagrams were projected/recorded in a way that was visible to everybody.	
The facilitation.	
The opportunity for open and extensive discussion.	
The use of causal diagrams.	
Simulation, using the quantitative model.	
Others,.....	

11. This will be the last time we ask. How would you allocate the investment of 10,000 monetary units now?

Formatted: English (United States)

Policy Area	***Amount in monetary units***
Energy Efficiency	
Monitoring	
Third Spaces	
Total	**10,000**

Please provide your name or any identifier that you will remember in the next workshops. It will help us track the development across workshops. If you want your replies to remain anonymous, please leave this blank.

Identifier:_____

Any other comments about the workshop and the HEW project:

Thank you again for your cooperation!

Pre-Test

Welcome!

Based on your experience with the built environment in London the Mayor has asked for your advice. Your task is to determine the allocation of an investment fund to improve the performance of the built environment in a mixed housing area of London. The Mayor has arrived at three different policy areas that he considers most important: energy efficiency, monitoring and third spaces.

- ✓ By investing in ***energy efficiency*** you are diversely supporting measures intended to reduce energy usage in buildings. This may take the form of subsidizing new technologies, promoting the benefits of retrofitting to encourage residents to invest and much more.

- ✓ Investments in ***monitoring*** are supportive of policies aimed towards collecting information about the performance of the built environment and implementing a spirit of accountability among architects, builders and developers.

- ✓ Investments in ***third spaces*** are directed towards adding or improving communal areas that can be accessed by residents. These areas are tertiary to homes (first space), offices (second space). Examples of third spaces are green spaces, pubs, cafés, recreational facilities and more.

Policy Area	*Amount in MU*
Energy Efficiency	
Monitoring	
Third Spaces	

Total	10,000

The mayor has **10,000 MU** or *monetary units* to invest across all three areas. Please consider *based on your experience*, and submit your recommendation below.

Post-Test

Now the Mayor has many new strategies to consider for allocating the 10,000 MU. Noticing your keen aptitude for this sort of work he has also asked for your **individual opinion.**

Policy Area	*Amount in MU*
Energy Efficiency	
Monitoring	
Third Spaces	
Total	10,000

Thank you!

Game Supplements

Log Sheets for the Game

TEAM NAME: _____ SIMULATION RUN #: _____

UNANIMOUS investment decision:

Policy Area	***Amount***
Energy Efficiency	
Monitoring	
Third Spaces	
Total	10,000

1. ***What outcome do you expect from these investments?***

2. ***Is the simulation outcome different? If so, what do you think is the cause of this?***

3. ***Other comments:***

Managing the discussion:

- **Remain neutral**: Do not take a position or give stakeholders advice on the investments.
- Be able to **address questions about model validity**. Address participants disagreeing with simulation outcomes.
    - Do so without explaining the dynamics of the model, save this for the second simulation round.
        - Remind them we are aiming to increase understanding and promote discussion.
        - Thank them for their criticisms, assure them they will have a chance to dig into the model assumptions and that this is how we can improve the model.
- Watch out for **talking heads** who keep the floor for extended periods. Don't be afraid to politely interject and allow others a chance to speak.
- Similarly, look out for people who have not spoken up. You don't need to force them to speak or address them directly but try to ensure they have opportunities to express their opinion.
- A key part of your role is to **keep discussion focused on the simulation environment**. Use some common sense here, for example:
    - **Good:** A stakeholder is talking about a personal experience to argue for their desired investment amounts.
    - **Bad:** Stakeholders debating the effectiveness of the Green Deal without making any connection to the simulations or investment decisions.

Please turn over for simulation instructions

Managing the simulation:

Part One: the first 3 simulations will be structured as follows, you will control the model:

- **Do not run any simulations if participants do not agree unanimously!!** Before each run, ask the group if everyone agrees with the decisions. If there is disagreement, do not run the simulation
- **Investments must add to 10,000**
- Record participant discussion on the provided log-sheets to the best of your ability. We want participants to think critically about their decisions, so do not run any simulations before discussing participant opinions on what will happen.

  **** But remember, your priority is to manage the group.** We will be recording the participants and the observer will also have a log sheet.

Part Two: Same as above for two simulations. Provide the participants with handouts of the CLD. **Stay neutral!** You can help them with the dynamics but don't give them your opinion on what is a high leverage input, etc.

Once the two structured simulations have been completed, participants can "play" with the environment.

- Briefly demonstrate how to modify other assumptions in the model, and allow them to If they have their own laptops, offer the URL.

Operating the model:

- You must "store" the run before moving the input sliders in order to create a comparative graph. (If you forget, no big deal, just reset the sliders to 0 and store the run.)
- Be careful not to reveal the CLD or S&F diagram during the first trial-driven simulation period.

Instructions Observer: Game Workshop

Key elements to look for:

Conflict: Describe conflict according to three domains: Task, interpersonal (relationship), and method (process). Conflict types can occur independently or all show up at once.

- ✓ Does it pertain to the *task* at hand? (i.e. discussions about the investments)
    - o This kind of conflict is what we are looking for, and should sound constructive. Disagreement is to be expected, but if negative language appears remember to look for other kinds of conflict.
- ✓ Is it of an *interpersonal* nature? (i.e. statements directed at people rather than the subject matter)
    - o Do you notice any *defensive* behavior? (i.e. participant takes an argument about the investments or behavior personally)
- ✓ Do they disagree with the *method* used? (i.e. distrust of the simulation environment)?

Power: Differences in seniority or decision making authority can affect the group process.

- ✓ Take inventory of the entities represented by your stakeholder group. (Policymakers, NGO, industry, community member etc.)
- ✓ Do any stakeholders use their position as an asset to support their arguments?
- ✓ Can you ascertain which stakeholders have more experience? Newcomers may be less assertive than those who have been involved in HEW project longer.
    - **Be watchful for *talking heads*, those who hold the speaking floor for long periods and/or interrupt other stakeholders often. Is this person more experienced, or representing a powerful group?
- ✓ Similarly, notice stakeholders who are quiet. Consider if power differences may be preventing them from speaking up. (e.g. Manager and employee in the same group etc.)

* IN ADDITION to your observations, you will also have a "log sheet" for investment decisions. Pay attention to the SD facilitator, and help record this information if they become too consumed with managing stakeholders.

** DON'T BE AFRAID TO ACT ON YOUR OBSERVATIONS! Although the SD Facilitator role is primarily focused on managing the group, be aware that they may fail to recognize when individuals dominate (or fail to join) discussions. Bring this this to the attention of the SD facilitator as subtly as possible.